

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|  |  |  |   |
|--|--|--|---|
| (51) International Patent Classification <sup>6</sup> :<br><b>C12N 15/11, 15/19, 15/24, 15/85</b>  |  | <b>A2</b>  | (11) International Publication Number: <b>WO 95/28482</b> |
|  |  | (43) International Publication Date: 26 October 1995 (26.10.95)  |   |
| (21) International Application Number: <b>PCT/US95/04477</b>   |  | (81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG). |   |
| (22) International Filing Date: 10 April 1995 (10.04.95)   |  |  |   |
| (30) Priority Data:<br>08/228,935 14 April 1994 (14.04.94) US<br>08/410,780 27 March 1995 (27.03.95) US  |  |  |   |
| (71) Applicant: <b>LIGAND PHARMACEUTICALS INCORPORATED (US/US); 9393 Towne Centre Drive, San Diego, CA 92121 (US).</b>   |  | Published<br><i>Without international search report and to be republished upon receipt of that report.</i>   |   |
| (72) Inventors: <b>SEIDEL, H., Marlin, Apartment H317, 5370 Toscana Way, San Diego, CA 92122 (US). LAMB, L. Peter; 7565 North Rim Court #291, San Diego, CA 92121 (US).</b>  |  |  |   |
| (74) Agents: <b>JURGENSEN, Thomas, E. et al; Ligand Pharmaceuticals Incorporated, 9393 Towne Centre Drive, San Diego, CA 92121 (US).</b>   |  |  |   |
| (54) Title: <b>DNA SPACER REGULATORY ELEMENTS RESPONSIVE TO CYTOKINES AND METHODS FOR THEIR USE</b>  |  |  |   |
| (57) Abstract  |  |  |   |
| <p>The present invention provides oligonucleotide sequences comprising DNA regulatory elements of the general sequence TTN<sub>x</sub>AA that bind activated transcriptional regulatory proteins in response to signaling molecules, such as cytokines. Further, the present invention also provides DNA constructs comprising the oligonucleotide sequences, cells transfected with the DNA constructs, and methods of using the DNA constructs and transfected cells to provide for the controlled expression of structural genes, for the detection and recovery of transcriptional regulatory proteins, and for measuring the ability of compounds to act as agonists and antagonists of gene transcription.</p> |  |  |   |

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                       |    |                          |
|----|--------------------------|----|---------------------------------------|----|--------------------------|
| AT | Austria                  | GB | United Kingdom                        | MR | Morocco                  |
| AU | Australia                | GE | Georgia                               | MW | Malawi                   |
| BB | Barbados                 | GN | Guinea                                | NE | Niger                    |
| BE | Belgium                  | GR | Greece                                | NL | Netherlands              |
| BF | Burkina Faso             | HU | Hungary                               | NO | Norway                   |
| BG | Bulgaria                 | IE | Ireland                               | NZ | New Zealand              |
| BJ | Benin                    | IT | Italy                                 | PL | Poland                   |
| BR | Brazil                   | JP | Japan                                 | PT | Portugal                 |
| BY | Belarus                  | KE | Kenya                                 | RO | Romania                  |
| CA | Canada                   | KG | Kyrgyzstan                            | RU | Russian Federation       |
| CF | Central African Republic | KP | Democratic People's Republic of Korea | SD | Sudan                    |
| CG | Congo                    | KR | Republic of Korea                     | SE | Sweden                   |
| CH | Switzerland              | KZ | Republic of Kazakhstan                | SI | Slovenia                 |
| CI | Cote d'Ivoire            | LA | Laos                                  | SK | Slovakia                 |
| CM | Cameroon                 | LI | Liechtenstein                         | SN | Senegal                  |
| CN | China                    | LK | Sri Lanka                             | TD | Chad                     |
| CS | Czechoslovakia           | LU | Luxembourg                            | TG | Togo                     |
| CZ | Czech Republic           | LV | Latvia                                | TJ | Tajikistan               |
| DE | Germany                  | MC | Monaco                                | TT | Trinidad and Tobago      |
| DK | Denmark                  | MD | Republic of Moldova                   | UA | Ukraine                  |
| ES | Spain                    | MG | Madagascar                            | US | United States of America |
| FI | Finland                  | ML | Mali                                  | UZ | Uzbekistan               |
| FR | France                   | MN | Mongolia                              | VN | Viet Nam                 |
| GA | Gabon                    |    |                                       |    |                          |

WO 95/28482

PCT/US95/04477

2

initiates a cascade of serine/threonine phosphorylations through MAP kinases leading to phosphorylation of DNA binding proteins, thereby changing their transcriptional modulatory activity. S.A. Moodie et al., 260 *Science*, 1658 (1993); C. A. Lange-Carter et al., 260 *Science*, 315 (1993); C. S. Hill et al., 73 *Cell* 395 (1993); H. Gille et al., 358 *Nature*, 414 (1992) and R. H. Chen et al., 12 *Mol. Cell. Biol.*, 915 (1992).

Despite these advances, the signal transduction pathways utilized by many growth factors and cytokines to alter gene expression remain unclear. Thus, although known second messengers have been implicated in signal transduction in response to some of these factors, their role in modulating gene expression remains speculative. Miyajama, 10 *Annu. Rev. Immunol.*, 295 and D.E. Levy and J.E. Darnell, 2 *New Biol.*, 923 (1990). This in turn raises the question of how ligand specific responses are elicited in such cellular systems. Ullrich, 61 *Cell*, 203; M.V. Chao, 68 *Cell*, 995 (1992) and Levy, 2 *New Biol.*, 923.

Progress in resolving these issues has been made recently in the interferon (IFN) system. IFNs  $\alpha$  and  $\beta$  (type I) act as a primary non-specific defense against viral infections. S. Patska and J.A. Langer, 56 *Annu. Rev. Biochem.*, 727 (1987). IFN $\gamma$  (type II) has anti-viral properties but also plays a major role in regulation of the immune response. *Id.* Type I and type II IFNs bind to distinct cell surface receptors and cause rapid alterations in gene expression. Aguet, 55 *Cell*, 273; Uze, 60 *Cell*, 225; and G.C. Sen and P. Lengyel, 267 *J. Biol. Chem.*, 5017 (1992). Specific sequence elements have been identified in the promoters of genes that respond to IFN $\alpha$ , termed interferon- $\alpha$  stimulated response elements (ISREs), that are both necessary and sufficient for regulation by IFN $\alpha$ . Sen, 267 *J. Biol. Chem.*, 5017. Specifically, activation of the IFN $\alpha$  receptors stimulates tyrosine phosphorylation of a family of proteins that serve as DNA binding proteins, and accordingly as transcription regulatory factors via the ISRE. C. Schindler et al., 257 *Science*, 809 (1992); K. Shuai et al., 258 *Science*, 1808 (1992) and M. J. Gutch et al., 89 *Proc. Natl. Acad. Sci. USA*, 11411 (1992). These DNA binding proteins, generically termed "signal transducers and activators of transcription" (STATs), assemble into a multimeric complex, translocate to the nucleus, and bind cis-

WO 95/28482

PCT/US95/04477

3

acting enhancer elements in the appropriate regulatory regions. D.E. Levy et al., 3 Genes Dev., 1362 (1989); and D.S. Kessler et al., 4 Genes Dev., 1753 (1990) and Z. Zhong et al., 264 Science, 95 (1994).

One example of an IFN $\alpha$ -induced ISRE binding protein complex is

- 5 ISGF3. T.C. Dale et al., 86 Proc. Natl. Acad. Sci., 1203 (1989) and X-Y. Fu et al., 87 Proc. Natl. Acad. Sci., 8555 (1990). ISGF3 is a complex of 4 binding proteins, called p48, p84 (STAT1 $\beta$ ), p91 (STAT1 $\alpha$ ) and p113 (STAT2). Recently, cDNAs encoding the proteins that constitute ISGF3 have been isolated and characterized. X-Y Fu et al., 89 Proc. Natl. Acad. Sci., 7840 (1992); C. Schindler et al., 89 Proc. Natl. Acad. Sci., 7836 (1992) and S.A. Veals et al., 12 Mol. Cell. Biol., 3315 (1992). p48 is the DNA binding component of ISGF3 and has homology to *myb*. Veals, 12 Mol. Cell. Biol., 3315. p84 and p91 are probably alternatively spliced products of the same gene and are related to p113. X-Y Fu, 89 Proc. Natl. Acad. Sci., 7840 and Schindler, 89 Proc. Natl. Acad. Sci., 7836. p84, p91 and p113 are novel proteins that contain SH2 and SH3 domains and are found in the cytoplasm of untreated cells. Schindler, 257 Science, 809 and X.Y. Fu, 70 Cell, 323-335 (1992). Thus, IFN $\alpha$  treatment of cells results in rapid tyrosine phosphorylation of p84, p91 and p113, causing them to associate and form a heteromeric complex with p48 to form ISGF3, which then translocates to the nucleus and binds to ISREs, stimulating transcription. Id.; Dale, 86 Proc. Natl. Acad. Sci., 1203 and Kessler, 4 Genes Dev., 1753.

Regulation in response to IFN $\gamma$  is conferred by a distinct sequence from the ISRE, the gamma activated sequence (GAS). T. Decker et al., 10 EMBO J., 927 (1991); K.D. Khan et al., 90 Proc. Natl. Acad. Sci., 6806 (1993) and D.J. Lew et al., 11 Mol. Cell. Biol., 182 (1991). Treatment of cells with IFN $\gamma$  results in tyrosine phosphorylation of p91 (STAT1 $\alpha$ ), which then translocates to the nucleus and binds to the GAS. Decker, 10 EMBO J., 927 and K. Shuai et al., 258 Science, 1808 (1992). Thus the specificity of binding of either IFN $\alpha$  or IFN $\gamma$  to their receptors is translated into a specific phosphorylation pattern within a related family of latent transcription factors (i.e. DNA binding proteins). This pattern of phosphorylation dictates the

WO 95/28482

PCT/US95/04477

4

association state of the proteins, which determines specificity of binding to either an ISRE or a GAS, and the subsequent transcriptional response.

- Yet another cytokine, Interleukin-6 (IL-6) plays a major role in the induction of the acute phase response in hepatocytes. The acute phase response is characterized by the dramatic transcriptional upregulation of a distinct set of genes, termed acute phase response genes. P.C. Heinrich et al, 265 Biochem. J., 621-636 (1990). Studies of the promoter regions of these genes have identified specific DNA sequences that are required for induction of acute phase response genes by IL-6. See D.R. Kunz et al., 17 Nuc. Acids Res., 1121-1138 (1989); M. Hattori et al., 87 Proc. Natl. Acad. Sci. USA, 2364-2368 (1990); K.A. Won and H. Baumann, 10 Mol. Cell. Biol., 3965-3978 (1990) and D.R. Wilson et al., 10 Mol. Cell. Biol., 6181-6191 (1990). These sequences are termed acute phase response elements (APREs). One type of APRE shows many similarities to the GAS elements that confers induction by IFN $\gamma$ . Yuan et al., 14 Mol. Cell. Biol., 1657-1668 (1994). Proteins that bind to this class of APREs have been characterized and purified. U. M. Wegenka et al., 13 Mol. Cell. Biol., 276-288 (1993); T. Ito et al., 17 Nuc. Acids Res., 9425-9435 (1989) and Hattori, 87 Proc. Natl. Acad. Sci. USA, 2364-2368. A cDNA clone encoding the IL-6-induced APRE-binding protein has been isolated (Zhong, 264 Science, 95; Akira et al., 77 Cell 63-71 (1994); Zhong, et al., 91 Proc. Natl. Acad. Sci., 4806-4810 (1994); Raz et al., 269 J. Biol. Chem., 24391-24395 (1994)), and was found to encode a protein that shows considerable homology to p91 (STAT1 $\alpha$ ). For this reason the protein is termed STAT3. Like STAT1 $\alpha$ , STAT3 is a latent transcription factor that is activated to bind DNA by rapid tyrosine phosphorylation.

- Interleukin-4 (IL-4) is a pleiotropic cytokine that elicits biological responses in a variety of both lymphoid and non-lymphoid cell types. IL-4 is a glycoprotein of approximately 19 kD produced primarily by the Th2 subset of activated T-cells. IL-4 has since been shown to play an important role in B-cell proliferation, the regulation of immunoglobulin expression, in T-cell regulation and in the growth and differentiation of hematopoietic precursor cells. IL-4 exerts its biological effects

WO 95/28482

PCT/US95/04477

5

through a specific high-affinity receptor on the surface of hematopoietic as well as certain non-hematopoietic cell lines. One chain of its receptor, the  $\gamma_c$  chain, is shared by the IL-2, IL-7, IL-9, IL-13 and IL-15 receptors. M. Kondo et al., 262 Science 1874 (1993), M. Noguchi et al., 262 Science 1877 (1993), S. Russell et al., 262 Science 1880 (1993), and M. Kondo et al., 263 Science 1453 (1994).

Binding of IL-4 to its receptor on the cell surface results in the activation of an intracellular tyrosine kinase and the rapid phosphorylation of several proteins on tyrosine. These initial events appear to be directly related to the immediate effects of IL-4 on target gene transcription. In particular, IL-4 up-regulates in responsive cell lines the expression of several cell-surface antigens including class II MHC, the low affinity Fc receptor for IgE (FcεRII, CD23), LFA-1 and LFA-3, CD40 and surface IgM. B. Aggarwal and J. U. Gutterman, Human Cytokines: Handbook for Basic Chemical Research, Blackwell Scientific Publications, Boston, MA (1992). Perhaps the most prominent role of IL-4 is in B-cell differentiation, where IL-4 acts as a "switch factor" promoting an Ig heavy chain class switch to IgE, the major mediator of Type I allergic reactions. W.E. Paul, 77 Blood 1859 (1991). Evidence that IL-4 operates through a STAT signal transduction system is based upon the observation that IL-4 rapidly activates in a variety of cell lines phosphotyrosine-containing protein complexes that bind to a GAS-like DNA sequence element. H. Kotanides and N. Reich 262 Science 1265 (1993) and C. Schindler et al., 13 EMBO J. 1350 (1994); P. Lamb et al., 83 Blood, 2063 (1994) and I. Köhler and E.P. Rieber, 23 Eur. J. Immunol., 3066 (1993). A STAT activated by IL-4 in THP-1 cells has been cloned recently (called STAT-IL-4 or STAT6) and is likely a constituent of the reported IL-4 induced complexes. J. Hou et al., 265 Science, 1701 (1994) and J.N. Ihle et al., 11 Trends in Genetics, 69 (1995).

Interleukin 13 (IL-13) is a pleiotropic cytokine that shares many of the biological activities of IL-4. G. Zurawski and J. E. de Vries, 15 Immunol. Today 19 (1994). IL-13 has roughly 30% sequence identity with IL-4 and exhibits IL-4-like activities on monocytes/macrophages and B-cells (A. Minty et al., 362, Nature 248 (1993) and A.N.J. McKenzie et al., 90 Proc. Natl. Acad. Sci. USA 3735 (1993).

WO 95/28482

PCT/US95/04477

6

- However, unlike IL-4, IL-13 has no effect on T-cells. The biological activity of IL-13 is mediated through binding to its specific high-affinity cell surface receptors consisting of an IL-13 binding subunit and one or more receptor components that are shared with the IL-4 receptor (the 'IL-4R' subunit and/or the  $\gamma c$  subunit). G. Aversa et al., 178 *J. Exp. Med.*, 2213 (1993). Evidence that IL-13, like IL-4, operates through a STAT signal transduction system is based upon the observation that IL-13 rapidly activates phosphotyrosine-containing protein complexes very similar to those induced by IL-4 that bind to a GAS-like DNA sequence element. I. Köhler et al., 345 *FEBS Lett.* 187 (1994).
- GM-CSF belongs to a group of growth factors termed colony stimulating factors which are involved in the survival, clonal expansion, and differentiation of hematopoietic progenitor cells. J. Gasson, 77 *Blood* 1131 (1991) and N.A. Nicola, 58 *Annu. Rev. Biochem.* 45 (1989). GM-CSF acts on a set of partially committed progenitor cells and causes them to divide and differentiate in the granulocyte-macrophage pathways. GM-CSF can also activate mature granulocytes and macrophages. In addition to effects on myelomonocytic lineages, GM-CSF can promote the proliferation of erythroid and megakaryocyte progenitor cells. GM-CSF, an 18-22 kD glycoprotein, is produced by a variety of cells, including T-cells, B-cells, macrophages, mast cells, endothelial cells and fibroblasts, in response to immune or inflammatory stimuli.
- GM-CSF exerts its effects by interacting with cell surface receptors on specific target cells. The receptor is composed of two chains, GM-CSF- $\alpha$  and GM-CSF- $\beta$ . L.S. Park et al., 89 *Proc. Natl. Acad. Sci.* 4295 (1992). The GM-CSF- $\alpha$  is specific to GM-CSF, while the GM-CSF- $\beta$  is identical to the  $\beta$  subunit of the IL-5 and IL-3 receptors. G. Goodall et al., 8 *Growth Factors* 87 (1993). Although neither GM-CSF $\alpha$  or GM-CSF $\beta$  have intrinsic kinase activity, GM-CSF treatment of cells results in rapid tyrosine phosphorylation of multiple proteins. Evidence that GM-CSF operates through a STAT signal transduction system is based upon the observation that GM-CSF rapidly activates in a variety of cell lines phosphotyrosine-containing protein complexes

WO 95/28482

PCT/US95/04477

7

that bind to a GAS-like DNA sequence element. A. C. Larner et al., 261 *Science* 1730 (1993) and P. Lamb et al., 83 *Blood* 2063 (1994). It has been reported that GM-CSF activates STAT5, which thus may be a constituent of the reported GM-CSF activated complexes. Ihle et al., 11 *Trends in Genetics*, 69 (1995).

- 5 Interleukin-3 (IL-3) is a pleiotropic cytokine produced primarily by activated T-cells. Its effects include stimulating the proliferation and differentiation of both pluripotent hematopoietic precursor cells as well as a wide variety of lineage committed cells Ihle, J.N. in *Peptide Growth Factors and their Receptors*, Springer-Verlag, New York (1991). The mature protein has an apparent molecular weight of 28,000, and binds to a cell surface receptor (IL-3R) that consists of at least two 10 polypeptide chains, IL-3R $\alpha$  and IL-3R $\beta$ . The IL-3R $\beta$  chain is also a component of the IL-5 and GM-CSF receptors, whereas the IL-3 $\alpha$  chain is unique to the IL-3R. Miyajama et al 82 *Blood* 1960, (1993). Binding of IL-3 to its receptor causes the activation of the tyrosine kinase JAK2 and the rapid tyrosine phosphorylation of a set of 15 cytoplasmic proteins. O. Silvennoinen et al., 90 *Proc. Natl. Acad. Sci.* 8429 (1993). A GAS-binding complex that contains a member of the STAT family can be detected in extracts from cells treated with IL-3. A.C. Larner et al., 261 *Science* 1730 (1993); J.N. Ihle et al., 19 *Trends Biochem. Sci.* 222 (1994). It has been reported that IL-3 activates STAT5, which thus may be a constituent of the reported IL-3-activated complexes. 20 J.N. Ihle et al., 11 *Trends in Genetics*, 69 (1995).

- Erythropoietin (Epo) is the major hormone responsible for the proliferation and maturation of red blood cell precursors. S.B. Krantz, 77 *Blood* 419 (1991). *In vitro* evidence indicates that it also plays a role in thrombocytopoiesis. An et al., 22 *Exp. Hemat.* 149 (1994). The protein, which has an apparent molecular weight 25 of 30,000, is produced mainly in the kidneys and is induced by conditions of tissue hypoxia. It acts by binding to a cell surface receptor (EpoR) that consists of a single polypeptide chain that is a member of the hematopoietin receptor family. A. D'Andrea et al 57 *Cell* 277 (1989). An early event following the binding of Epo to EpoR is the activation of the tyrosine kinase JAK2, which associates non-covalently with the 30 cytoplasmic domain of the receptor chain. B. Witthuhn et al, 74 *Cell* 227. Activation



WO 95/28482

PCT/US95/04477

8

of JAK2 by Epo is correlated with induction of tyrosine phosphorylation of the EpoR and cytoplasmic proteins. Epo treatment of cells also results in the rapid induction of a GAS-binding activity that contains as yet unidentified STAT proteins, that are thought to contribute to Epo-induced changes in gene expression. P. Lamb et al., 83 Blood 2063 (1994); Finbloom et al., 14 Mol. Cell Biol. 2113 (1994). It has been reported that Epo  
 5 activates STAT5, which thus may be a constituent of the reported Epo-activated complexes. J.N. Ihle et al., 11 Trends in Genetics, 69 (1995).

G-CSF is a pleiotropic cytokine best known for its specific effects on the proliferation, differentiation, and activation of hematopoietic cells of the neutrophilic  
 10 granulocyte lineage. G-CSF has also been reported to have chemotactic activity for human granulocytes and monocytes as well as for mesenchymal cells including fibroblasts, smooth muscle cells and myofibroblasts. These in vitro functions reflect the potential in vivo roles for G-CSF in the maintenance of steady state hemopoiesis, defense against infection, inflammation and repair. When G-CSF was administered to  
 15 various animal models, an elevation of circulating neutrophils has been observed. G-CSF is now used clinically in patients that have granulopenia as a result of receiving chemotherapy or receiving immunosuppressive agents after organ transplantation. M.A.S. Moore, 9 Annu. Rev. Immunol. 159 (1991), N.A. Nicola, 58 Annu. Rev. Biochem. 45 (1989), and E. Pimentel, (1994) in Handbook of Growth Factors, Vol III,  
 20 E. Pimentel, ed., CRC Press, Boca Raton, p. 177.

G-CSF exerts its biological activity through binding to G-CSFR. The receptor for G-CSF (G-CSFR) is a member of the type I cytokine receptor superfamily that lacks a kinase domain and appears to consist of a single polypeptide chain. Dimerization of two G-CSFR chains forms a high affinity binding site for G-CSF.  
 25 Among the various hematopoietin receptor superfamily members, G-CSFR is most closely related to gp130, the signal-transducing component of the IL-6, oncostatin M, and leukemia inhibitory factor receptors. Recent studies have demonstrated that in myeloid leukemia cell lines, G-CSF treatment results in rapid tyrosine phosphorylation of G-CSFR, JAK1 and JAK2 tyrosine kinases and the members of the STAT family of

WO 95/28482

PCT/US95/04477

9

*activated sequence*

transcription factors. S.E. Nicholson et al. 91 Proc. Natl. Acad. Sci. USA 2985 (1994) and S.S. Tian et al., 84 Blood 1760 (1994).

Although many cytokines activate STAT proteins that can bind to similar sequences (GAS/APREs), they regulate distinct sets of genes. This suggests that there is specificity with respect to the response elements in some of these genes, such that they respond only to one of these cytokines. It is therefore possible that distinct classes of GAS-like sequences exist that are selective in their ability to respond to various cytokines. Accordingly, specific DNA sequences that show selectivity with respect to their ability to bind cytokine-activated proteins, including STAT proteins, would be useful tools allowing the responses mediated by different cytokine-activated DNA-binding proteins to be assayed selectively.

The disclosures of the above-cited references are hereby incorporated by reference in their entirety.

15

#### Summary of the Invention

The present invention is directed to oligonucleotide sequences comprising DNA regulatory elements that bind, either directly or indirectly, to activated transcriptional regulatory proteins, preferably STAT proteins, in response to signaling molecules, including cytokines such as interferon gamma (IFN $\gamma$ ), interleukin 3 (IL-3), interleukin 4 (IL-4), interleukin 6 (IL-6), interleukin 13 (IL-13), erythropoietin (Epo), granulocyte colony stimulating factor (G-CSF) and granulocyte-macrophage colony stimulating factor (GM-CSF). Surprisingly, a spacing of 4, 5, 6 or 7 nucleotides between each half-site of the regulatory elements results in discrete regulatory elements that bind a variety of activated transcriptional regulatory proteins and/or specifically bind a single type or class of activated transcriptional regulatory protein. Accordingly, the regulatory elements of the present invention can be used in transcriptional assays to discover selective agonists or antagonists of a signaling molecule, such as IL-4, and its cognate transcriptional regulatory protein, STAT6.

In particular, the present invention provides an oligonucleotide sequence comprising a regulatory element of the nucleotide sequence TTN<sub>x</sub>AA, wherein N is

30

WO 95/28482

PCT/US95/04477

DNA SPACER REGULATORY ELEMENTS  
RESPONSIVE TO CYTOKINES  
AND METHODS FOR THEIR USE

5

Field of the Invention

This invention relates to oligonucleotide sequences that bind regulatory proteins that affect transcription in response to various molecules, such as cytokines, to DNA constructs comprising the oligonucleotide sequences, cells transfected with the DNA constructs, and to methods of using the same to provide for the controlled expression of heterologous genes, for the detection and recovery of new regulatory proteins, and for measuring the ability of compounds to act as agonist and antagonists of gene transcription.

A 10

15

Background of the Invention

In many cellular systems, extracellular signaling molecules, such as polypeptide ligands, bind to receptors on the surface of the cells, thereby triggering an intracellular signaling pathway that ultimately regulates gene transcription within the cells. For example, cytokines and growth factors, which comprise a large and diverse family of soluble polypeptides that control the growth, differentiation and function of mammalian cells, bind to specific cell surface receptors, that in some way transduce signals that elicit a specific phenotypic response. A. Miyajima et al., 10 Annu. Rev. Immunol., 295 (1992); M. Aguet et al., 55 Cell, 273 (1988); T. Kishimoto et al., 258 Science, 593 (1992) and A. Ullrich and J. Schlessinger, 61 Cell, 203 (1990). Abundant evidence shows that changes in the transcription rate of specific genes are an important component of this response. This is thought to be a consequence of alterations in the amount or the activity of specific DNA-binding proteins.

In some instances, progress has been made in defining the pathway that leads from a receptor-ligand interaction at the cell surface to changes in the activity of such DNA binding proteins or other nuclear proteins. Ulrich, 61 Cell 203. In this regard, a common response in surface receptor signaling pathways involves the activation of Ras. L.S. Mulcahy et al., 313 Nature, 241 (1985). Activated Ras then

30

independently selected from the nucleotides A, T, C or G and x is 4, 5, 6 or 7 and wherein the regulatory element binds to an activated transcriptional regulatory protein in response to a signaling molecule. More preferably, the regulatory elements comprise nucleotide sequences selected from the group consisting of  $\text{TTCN}_y\text{GAA}$ ,  $\text{TTAN}_y\text{GAA}$ ,

- 5  $\text{TTAN}_y\text{TAA}$ ,  $\text{TTTN}_y\text{GAA}$  and  $\text{TTTN}_y\text{TAA}$ , where N has the same meaning given above, and y is 2, 3, 4 or 5. These oligonucleotide sequences can be double stranded, including their complements.

The present invention also provides a DNA construct comprising the oligonucleotide sequence described above operably linked to a promoter, which  
10 promoter is operably linked to a heterologous gene, wherein the DNA construct is linked in such a manner that the heterologous gene is under the transcriptional control of the oligonucleotide sequence and promoter. Also provided is a host cell transfected with this DNA construct.

- The present invention also provides a method for the controlled  
15 expression of a heterologous gene of interest comprising culturing the transfected host cells containing an appropriate transcriptional regulatory protein(s) in the presence of a signaling molecule. Preferably, the signaling molecule in this method comprises a cytokine and the transcriptional regulatory protein comprises a STAT protein.

- The present invention further provides a method for detecting the  
20 presence of a transcriptional regulatory protein, such as a novel STAT protein, in a sample comprising contacting the sample with an oligonucleotide sequence as described above under conditions where the transcriptional regulatory protein is activated and binds with the oligonucleotide sequence to form a complex, and detecting the presence of the complex in the sample. Thereafter, the complex can be separated  
25 from the sample, and the transcriptional regulatory protein isolated from the regulatory element.

Further, the present invention provides a method for measuring the ability of a compound to act as an agonist of gene transcription comprising (a) contacting the compound with a transfected host cell as described above under

*inhibition  
of the binding  
of the oligonucleotide  
to the STAT protein*

#20

WO 95/28482

PCT/US95/04477

11

conditions in which the heterologous gene is capable of being expressed in response to the compound, and (b) comparing the level of gene expression in step (a) with the level of gene expression from the host cell in the absence of the compound. Alternatively, the present invention also provides a method for measuring the ability of a compound to act as an antagonist of gene transcription comprising (a) contacting the compound with a transfected host cell as described above in the presence of a predetermined amount of a signaling molecule under conditions in which the heterologous gene is capable of being expressed in response to the signaling molecule, and (b) comparing the level of gene expression in step (a) with the level of gene expression from the host cell in the presence of the signaling molecule, but the absence of the compound. In both these methods, the heterologous gene may be any appropriate reporter gene such as the heterologous gene for luciferase, chloramphenicol acetyl transferase, green fluorescent protein or  $\beta$ -galactosidase.

The present invention also provides an IL-4 and IL-13 selective oligonucleotide sequence comprising ANTTCNNNNGAANA (SEQ ID NO. 152) or its complement TTTTCNNNNGAANT (SEQ. ID NO. 153), wherein N is independently selected from A, T, C or G, and wherein the oligonucleotide sequence, when operably linked to a promoter and a heterologous gene, binds to an activated transcriptional regulatory protein comprising STAT6 protein activated by IL-4 and/or IL-13, and transcriptionally modulates the heterologous gene.

These and various other advantages and features of novelty which characterize the invention are pointed out with particularity in the claims annexed hereto and forming a part hereof. However, for a better understanding of the invention, its advantages, and objects obtained by its use, reference should be had to the accompanying drawings and descriptive matter, in which there is illustrated and described preferred embodiments of the invention.

#### Definitions

For the purposes of this invention:

"Oligonucleotide" or "DNA" molecule or sequence refers to a molecule comprised of the deoxyribonucleotides adenine (A), guanine (G), thymine (T) and/or cytosine (C), in either single-stranded form or a double-stranded helix, and comprises or includes a "regulatory element" according to the present invention, as that term is defined herein. The exact size, strandedness and orientation (i.e. 3' to 5', or 5' to 3') will depend upon many factors, which, in turn, depend upon the ultimate function and use of the oligonucleotides of the present invention. Thus, the term "oligonucleotide" or "DNA" includes double-stranded DNA found in linear DNA molecules or fragments, viruses, plasmids, vectors, chromosomes or synthetically derived DNA. As used herein, particular double-stranded DNA sequences may be described according to the normal convention of giving only the sequence in the 5' to 3' direction.

"Regulatory element" refers to a deoxyribonucleotide sequence comprising the whole, or a portion of, an oligonucleotide sequence to which an activated transcriptional regulatory protein, or a complex comprising one or more activated transcriptional regulatory proteins, binds so as to transcriptionally modulate the expression of an associated gene or genes, including heterologous genes.

"Signaling molecule" refers to an extracellular polypeptide, oligosaccharide or other non-peptidyl molecule, in either a free or bound form, that interacts with a receptor at or near the surface of a cell. This interaction in turn triggers an intracellular pathway which includes the activation of one or more transcriptional regulatory proteins that bind to a regulatory element, thereby transcriptionally modulating the expression of an associated gene or genes. As used herein, "signaling molecule" includes naturally occurring molecules, such as cytokines, peptidyl and non-peptidyl hormones, antibodies, cell-surface antigens, or synthetic mimics of any of these signaling molecules, or synthetic molecules that mimic the action of any of these signaling molecules.

"Cytokines" refer to a diverse grouping of soluble polypeptides, including growth factors and hormones, that control the growth, differentiation and function of cells in such a manner as to ultimately elicit a phenotypic response in an organism. Preferred cytokines useful with the regulatory elements and associated

#30

WO 95/28482

PCT/US95/04477

13

methods of the present invention include IFN $\gamma$ , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-11, IL-12, IL-13, IL-15, GM-CSF, Epo, Tpo, growth hormone, prolactin, Oncostatin M, G-CSF, LIF, EGF, CNTF and PDGF.

"Transcriptional regulatory protein" refers to cytoplasmic or nuclear proteins that, when activated, bind the regulatory elements/oligonucleotide sequences of the present invention either directly, or indirectly through a complex of transcriptional regulatory proteins or other adapter proteins, to transcriptionally modulate the activity of an associated gene or genes. Thus, transcriptional regulatory proteins can bind directly to the DNA regulatory elements of the present invention, or can bind indirectly to the regulatory elements by binding to another protein, which in turn binds to or is bound to a DNA regulatory element of the present invention. See e.g., S.A. Veals et al., 13 Molec. Cell Biol., 196-206 (1993). As used herein, transcriptional regulatory proteins, include, but are not limited to, those proteins referred to in the art as STAT proteins (Z. Zhong et al., 264 Science, 95) STF proteins (C. Schindler et al., 13 EMBO J., 1350 (1994)), Mammary Gland-Specific Nuclear Factor (M. Schmidt-Ney et al., 6 Mol. Endocrinol., 1988 (1992) and H. Wakao et al., 267 J. Biol. Chem., 16365 (1992)), APRF (Wegenka, 13 Mol. Cell Bio., 276), GHIF (Mayer, 269 J. Biol. Chem., 4701), GHSF and EPOSF (Finbloom, 14 Mol. Cell Bio., 2113), as well as to all substantially homologous analogs and allelic variations thereof.

"Transcriptionally modulate the expression of an associated gene or genes" means to change the rate of transcription of such gene or genes.

"STAT protein" refers to those transcriptional regulatory proteins designated as "Signal Transducers and Activators of Transcription" (STAT) by Dr. J.E. Darnell of Rockefeller University. See Zhong, 264 Science 95. As used herein, STAT proteins include the p91 (STAT1 $\alpha$ ), p84 (STAT1 $\beta$ ), p113 (STAT2) proteins and the STAT-associated p48 family of proteins. S.A. Veals et al., 12 Mol. Cell Biol., 3315 (1992). Further, STAT proteins also include a binding protein designated as STAT3 (Zhong, 264 Science 95), and a binding protein designated as STAT4 (Id.). In addition, MGF is now renamed STAT5 (Gouilleux et al., 13 EMBO J., 4361-4369 (1994)) and STAT-IL-4 (or STAT6) has recently been cloned. Hou et al., 265 Science, 730 (1994)

and J.N. Ihle et al., 11 Trends in Genetics, 69 (1995). Also included are substantially homologous analogs and allelic variations of all of the above STAT proteins.

"Activate", "activated", "activation" or derivatives thereof, means that one or more transcriptional regulatory proteins within a cell are modified post-translationally, or are constitutively active, such that they can bind directly or indirectly to DNA regulatory elements/oligonucleotide sequences of the present invention in a sequence-specific manner. This modification will typically comprise phosphorylation of the transcriptional regulatory proteins via a variety of mechanisms, including, but not limited to activation by various protein kinases. See, e.g., (Shuai, 258 Science, 1808 and P. Cohen, 17 TIBS, 408 (1992)).

"DNA construct" refers to any genetic element, including, but not limited to, plasmids, vectors, chromosomes and viruses, that incorporate the oligonucleotide sequences of the present invention. For example, the DNA construct can be a vector comprising a promoter that is operably linked to an oligonucleotide sequence of the present invention, which is in turn, operably linked to a heterologous gene, such as the gene for the luciferase reporter molecule.

"Promoter" refers to a DNA regulatory region capable of binding directly or indirectly to RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of the present invention, the promoter is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter will be found a transcription initiation site (conveniently defined by mapping with S1 nuclease), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CCAT" boxes. Prokaryotic promoters contain Shine-Dalgarno sequences in addition to the -10 and -35 consensus sequences.

"Gene" refers to a nucleic acid molecule, the sequence of which includes all the information required for the normal regulated production of a particular protein. A "heterologous" region of a DNA construct (i.e. a heterologous gene) is an identifiable



WO 95/28482

PCT/US95/04477

15

P50

segment of DNA within a larger DNA construct that is not found in association with the other genetic components of the construct in nature. Thus, when the heterologous gene encodes a mammalian gene, the gene will usually be flanked by a promoter that does not flank the structural genomic DNA in the genome of the source organism.

5 A promoter of a DNA construct, including an oligonucleotide sequence according to the present invention, is "operably linked" to a heterologous gene when the presence of the promoter influences transcription from the heterologous gene, including genes for reporter sequences such as luciferase, chloramphenicol acetyl transferase,  $\beta$ -galactosidase, green fluorescent protein and secreted placental alkaline  
10 phosphatase. Operably linked sequences may also include two segments that are transcribed onto the same RNA transcript. Thus, two sequences, such as a promoter and a "reporter sequence" are operably linked if transcription commencing in the promoter will produce an RNA transcript of the reporter sequence. In order to be "operably linked" it is not necessary that two sequences be immediately adjacent to one  
15 another.

A host cell has been "transfected" by exogenous or heterologous DNA (e.g. a DNA construct) when such DNA has been introduced inside the cell. The transfecting DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes, yeast, and mammalian cells for  
20 example, the transfecting DNA may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transfected cell is one in which the transfecting DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population  
25 of daughter cells containing the transfecting DNA.

"Host cell" refers to a cell line that expresses, either normally or after transfection of the requisite cDNAs, the relevant receptor components for a given signaling molecule, signaling (e.g., kinase) proteins, transcriptional regulatory proteins, and accessory factors such that, upon binding of the signaling molecule to the cell  
30 surface, transcriptional regulatory protein-mediated gene transcription is affected.

WO 95/28482

PCT/US95/04477

16

Preferably, the host cell line is responsive to cytokines, such that the host cell line expresses, either normally or after transfection of the requisite cDNAs, the relevant cytokine receptor components, JAK proteins, STAT proteins, and accessory factors such that, upon cytokine binding to the cell surface, STAT-mediated gene transcription is affected.

#### Brief Description of the Drawings

The invention may be further illustrated by reference to the accompanying Drawings wherein:

FIG. 1 is a schematic depiction of the reporter constructs that can be used to assess the transcriptional activity of transcriptional regulatory proteins that bind to the regulatory elements/oligonucleotide sequences of the present invention. Figure 1A depicts a general reporter construct that contains a regulatory element (single or multimerized) upstream of an operational promoter sequence, the two of which are operably linked to a downstream heterologous reporter gene sequence. Figure 1B depicts the more specific reporter constructs used in the experiments described in the Examples herein and summarized in Tables 7-10 and 15-16. Thus, the reporter constructs used contain a (multimerized) regulatory element located upstream of the Herpes Simplex virus thymidine kinase -35 to +10 (with respect to the cap site) promoter sequence, the two of which are operably linked to a downstream firefly luciferase reporter gene sequence; and

FIG. 2 is a reproduction of Electrophoretic Mobility Shift Assay (EMSA) autoradiograms that show the binding patterns of transcriptional regulatory protein-DNA binding complexes activated by IFN $\gamma$ , IL-6, GM-CSF and IL-4. The EMSA's were performed as described in Example 1 herein. The radiolabeled, double-stranded oligonucleotide probes utilized in the EMSAs of FIGS. 2A-2D were made by annealing the oligonucleotides of SEQ ID NOs 66 and 67 (FIG. 2A), 76 and 77 (FIG. 2B), 78 and 79 (FIG. 2C) and 80 and 81 (FIG. 2D).

WO 95/28482

PCT/US95/04477

17

Detailed Description of  
Embodiments of the Invention

The present inventors have discovered a series of DNA regulatory elements (i.e. response elements) that in response to various signaling molecules, bind, either directly or indirectly, to activated transcriptional regulatory proteins, and accordingly, transcriptionally modulate the expression of one or more genes operably linked with such regulatory elements. In this regard, the inventors have surprisingly discovered that a change in the spacing of 4, 5, 6 or 7 nucleotides within the regulatory elements can transform an element from one that binds with a variety of cytokine-induced, activated transcriptional regulatory proteins to a regulatory element that selectively binds just one type or class of cytokine-induced, activated transcriptional regulatory proteins. For example, in the sequence TTN<sub>x</sub>AA, a spacing of four spacer nucleotides (i.e. x = 4) results in a regulatory element that is selective for an activated STAT3 transcriptional regulatory protein (e.g., induced by IL-6, IL-10, IL-11, LIF, EGF, PDGF or G-CSF), while a spacing of six or seven spacer nucleotides yields a regulatory element that is selective for the activated transcriptional regulatory protein(s) induced by the IL-4, IL-7, IL-9, IL-13 and IL-15 cytokines. However, utilization of five spacer nucleotides results in a regulatory element that binds a variety of different cytokine-induced transcriptional regulatory proteins, including the STAT1 $\alpha$  and STAT3 proteins, as well as those transcriptional regulatory protein(s) that are induced by the IL-2, IL-3, IL-4, IL-5, IL-7, IL-9, IL-10, IL-11, IL-12, IL-13, IL-15, Epo, Tpo, growth hormone, prolactin, G-CSF and GM-CSF cytokines (e.g., STAT5 and/or STAT6).

The regulatory elements according to the present invention are selected from the nucleotide sequence TTN<sub>x</sub>AA, wherein N is a spacer nucleotide independently selected from A, T, C or G and x is 4, 5, 6 or 7. More preferably, the regulatory elements comprise nucleotide sequences selected from the group consisting of TTCN<sub>7</sub>GAA, TTAN<sub>7</sub>GAA, TTAN<sub>7</sub>TAA, TTTN<sub>7</sub>GAA and TTTN<sub>7</sub>TAA, where N

WO 95/28482

PCT/US95/04477

18

has the same meaning given above, and y is 2, 3, 4 or 5. The preferred nucleotide sequences according to the present invention include:

- |    |                            |                                |
|----|----------------------------|--------------------------------|
|    | TTNNNNAA (SEQ ID NO:1)     | TTNNNNNAA (SEQ ID NO:2)        |
| 5  | TTNNNNNAA (SEQ ID NO:3)    | TTNNNNNNNAA (SEQ ID NO:4)      |
|    | TTNNNGAA (SEQ ID NO:5)     | TTNNNGAA (SEQ ID NO:6)         |
|    | TTNNNNNGAA (SEQ ID NO:7)   | TTANNNGAA (SEQ ID NO:8)        |
|    | TTANNNGAA (SEQ ID NO:9)    | TTANNNGAA (SEQ ID NO:10)       |
|    | TTANNNNNGAA (SEQ ID NO:11) | TTANNNTTA (SEQ ID NO:12)       |
| 10 | TTANNNTTAA (SEQ ID NO:13)  | TTANNNNNTAA (SEQ ID NO:14)     |
|    | TTNNNGAA (SEQ ID NO:15)    | TTNNNGAA (SEQ ID NO:16)        |
|    | TTNNNNNGAA (SEQ ID NO:17)  | TTNNNNNGAA (SEQ ID NO:18)      |
|    | TTNNNTTAA (SEQ ID NO:19)   | TTNNNNNTAA (SEQ ID NO:20), and |
|    | TTNNNNNTTAA (SEQ ID NO:21) |                                |

- 15 where each N is independently selected from A, T, C, or G.

Especially preferred regulatory element sequences according to the present invention include: TTCCCGAA (SEQ ID NO. 22), TTCCCGAA (SEQ ID NO. 23), TTCCCGGAA (SEQ ID NO. 24), TTCCCGAA (SEQ ID NO. 25),

20 TTCCCGGAA (SEQ ID NO. 26), TTCCCGGAA (SEQ ID NO. 27), TTCTGGAA (SEQ ID NO. 28), TTCTTGGAA (SEQ ID NO. 29), TTCCAGAA (SEQ ID NO. 30), TTCCAGAA (SEQ ID NO. 31), TTCCCGAGAA (SEQ ID NO. 32), TTCTTGGAA (SEQ ID NO. 33), TTCTTGGAA (SEQ ID NO. 34), TTCTCAGAA (SEQ ID NO. 35), TTCTCCAGAA (SEQ ID NO. 36), TTACCGTAA (SEQ ID NO. 37),

25 TTACCCGTAA (SEQ ID NO. 38), TTACCCGTAA (SEQ ID NO. 39), TTCCCGTAA (SEQ ID NO. 40), TTCCCGTAA (SEQ ID NO. 41), TTCCCGGTAA (SEQ ID NO. 42), TTCTGTAA (SEQ ID NO. 43), TTCTCGTAA (SEQ ID NO. 44), TTCTCGTAA (SEQ ID NO. 45), TTCTCCGTAA (SEQ ID NO. 46), TTCCAGAA (SEQ ID NO. 47), TTCCCGTAA (SEQ ID NO. 48),

30 TTCCAGGAA (SEQ ID NO. 49), TTCTAAGAA (SEQ ID NO. 50), TTCTAAGAA (SEQ ID NO. 51), TTCCCGGAA (SEQ ID NO. 52), TTCTAAGAA (SEQ ID NO. 53).

WO 95/28482

PCT/US95/04477

19

In this regard, the regulatory elements of SEQ ID NOs 22, 25, 30 and 43 comprise STAT3 protein selective regulatory elements, SEQ ID NOs 37, 40 and 44 comprise STAT1 $\alpha$  and STAT3 selective regulatory elements, SEQ ID NOs 23, 26, 28, 31, 51, 52 and 53 comprise regulatory elements that bind a variety of activated transcriptional regulatory proteins, SEQ ID NOs 33, 35, 38, 41 and 45 comprise regulatory elements that bind GM-CSF, IL-2 and IL-4 activated transcriptional regulatory proteins, and SEQ ID NOs 24, 27, 29, 32, 34, 36, 39, 42, 46, 47, 48, 49 and 50 comprise regulatory elements that selectively bind activated transcriptional regulatory protein(s) induced by IL-4 and IL-13. Furthermore, these regulatory elements alone, or with additional flanking nucleotide sequences, form the oligonucleotide sequences according to the present invention. In this regard, it is preferable that such oligonucleotide sequences comprise between 8 and 200 nucleotides, including the regulatory elements of the present invention. However, sequences in excess of 200 nucleotides that contain the regulatory elements of the present invention, that are capable of binding activated transcriptional regulatory proteins, and of transcriptionally modulating the expression of one or more genes thereby, are also considered to be within the scope of the present invention. In this regard, it will be understood by those skilled in the art that, while the observed spacing preferences are an important determinant of selective binding and transcriptional activity, the particular sequences of the spacer nucleotides as well as the particular sequences of the nucleotides flanking the core elements also influence binding and transactivation.

The oligonucleotide sequences of the present invention can also comprise multimers of two or more "units" of the basic regulatory elements. In this regard, such multimer oligonucleotide sequences can, as a practical matter, contain from about 2 to 15 units of the same or different regulatory elements according to the present invention. However, theoretically, there is no limit to the number of regulatory elements within such a multimer oligonucleotide sequence. Such multimeric oligonucleotide sequences are useful as probes for detecting, isolating and/or purifying transcriptional regulatory proteins. Further, when used in a DNA construct, including a

WO 95/28482

PCT/US95/04477

20

promoter and heterologous gene, according to the present invention, a multimer of the regulatory elements can enhance the expression of the gene from the DNA construct in response to various cytokines or other signaling molecules. 470

A variety of signaling molecules activate transcriptional regulatory proteins that bind directly or indirectly to the regulatory elements/oligonucleotide sequences of the present invention. Nonlimiting examples of such signaling molecules include polypeptides such as cytokines and antibodies, and cell-surface antigens, oligosaccharides typically found at or near the surface of cell, non-peptidyl molecules such as TUBag4 (P. Constant et al., 264 *Science*, 267 (1994)) and synthetic mimics of any of these molecules, in both their free and bound forms. Thus, the present invention includes cell to cell or cell to substrate transcriptional regulatory protein activation via signaling molecules bound to or near the surface of a cell or other substrate.

Preferably, the signaling molecules according to the present invention comprise cytokines that activate transcriptional regulatory proteins that bind to the regulatory elements/oligonucleotide sequences of the present invention. Examples of such cytokines include, but are not limited to, Interleukins 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13 and 15 (IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-11, IL-12, IL-13 and IL-15), granulocyte-macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), colony stimulating factor 1 (CSF-1), interferons alpha, beta, and gamma (IFN $\alpha$ , IFN $\beta$ , IFN $\gamma$ ), epidermal growth factor (EGF), platelet-derived growth factor (PDGF), leukemia inhibitory factor (LIF), Oncostatin M, nerve growth factor (NGF), ciliary neurotrophic factor (CNTF), brain-derived neurotrophic factor (BDNF), erythropoietin (Epo), thrombopoietin (Tpo), growth hormone and prolactin. Particularly preferred cytokines according to the present invention include, but are not limited to, IFN $\alpha$ , IFN $\gamma$ , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-11, IL-12, IL-13, IL-15, G-CSF, GM-CSF, Epo, Tpo, Oncostatin M, G-CSF, LIF, EGF, PDGF and CNTF.

The regulatory elements and/or oligonucleotide sequences of the present invention will also prove useful in detecting, isolating and purifying new transcriptional regulatory proteins that display binding specificity to the regulatory elements/ 480

WO 95/28482

PCT/US95/04477

21

oligonucleotide sequences of the present invention. Further, it is contemplated that these regulatory elements/oligonucleotide sequences will prove particularly useful in the discovery of novel STAT proteins or STAT-related transcriptional regulatory proteins. In this regard, detection of such novel transcriptional regulatory proteins can be accomplished with the following technique. Cells, such as HepG2 cells, are treated with an appropriate cytokine, for example, with IFN $\gamma$  for 15 minutes to induce the activation of one or more transcriptional regulatory proteins. Extracts of the nucleus and cytoplasm of these cells are then prepared using conventional methods and tested for binding to the regulatory elements/oligonucleotide sequences by an electrophoretic mobility shift assay, in comparison with untreated cells that will show little or no specific binding as described in Levy, 3 *Genes Dev.*, 1362 and Kessler, 4 *Genes Dev.*, 1753, the disclosures of which are herein incorporated by reference. Furthermore, DNA regulatory element binding activity may also be stimulated *in vitro* by treating a cytoplasmic extract, supplemented with cell membranes, with a signaling molecule, such as a cytokine.

If an antibody specific for a transcriptional regulatory protein is available, it can be used to specifically interfere with the binding of the regulatory element of the present invention to the activated transcriptional regulatory protein, thereby assisting in the identification of the transcriptional regulatory protein. Furthermore, an unknown transcriptional regulatory protein identified or purified using the regulatory elements/oligonucleotide sequences of the present invention can be used to immunize animals to prepare an antibody specific for the transcriptional regulatory protein using methods well known in the art. See, e.g., E. Harlow, et al., *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1988), the disclosure of which is herein incorporated by reference.

Thus, the regulatory elements/nucleotide sequences of the present invention thus can serve as a "probe", similar to those used in a variety of nucleic acid detection systems well known in the art, except that the probes of the present invention are used to detect proteins, rather than a nucleic acid sequences, which specifically bind to the regulatory elements/oligonucleotide sequences of the present invention.

WO 95/28482

PCT/US95/04477

22

The sensitivity of such a nucleic acid detection assay can be increased by altering the manner in which a signal is detected by an observer. For example, assay sensitivity can be increased through the use of labeled oligonucleotide sequences using a wide variety of detectable labels, including, without limitation, enzyme labels, radioisotopic labels, fluorescent labels, and modified bases. See, e.g., U.S. Patent Nos. 4,581,333, 4,358,535, 4,446,237, 4,582,789, and 4,563,417, as well as European Patent Application Nos. EP 144914 and EP 119448, the disclosures of which are herein incorporated by reference. Thus, DNA probes according to the present invention preferably include the regulatory elements alone, or as part of a longer oligonucleotide sequence of the present invention, labeled with a detectable label, such as a radioisotope, an enzyme, a fluorescent label, a chemical label, or a modified base.

Thus, the present invention provides a method for detecting the presence of novel transcriptional regulatory proteins in a sample. Such samples are preferably biological samples, including, but not limited to, cells, cell culture supernatant, cell or tissue extracts, or particular fractions thereof, and other biological fluids such as blood, sera, urine, saliva, etc. Binding of the probe containing the regulatory elements/oligonucleotide sequences of the present invention to a transcriptional regulatory protein in the sample may be detected by any appropriate means known in the art. For example, direct or indirect, or competitive binding assays may be used. In such assays, association of the labeled probe with the proteinaceous material of the sample is then detected. In a preferred embodiment, the oligonucleotide sequence is modified by the incorporation of a radioactively labeled nucleotide.

Once detected, the novel transcriptional regulatory protein can be separated and purified from the probe-protein complex by any of a variety of techniques well known to those of skill in the art. For example, such isolation and purification can be based on affinity chromatography, which relies on the interaction of the protein to be purified with an immobilized ligand. In the present invention, a regulatory element and/or oligonucleotide sequence of the present invention immobilized on a support would serve as the immobilized ligand, which in turn would be used to isolate and purify a novel transcriptional regulatory protein from a sample.



WO 95/28482

PCT/US95/04477

23

In a preferred embodiment, the regulatory element/oligonucleotide sequence of the present invention is immobilized on a solid support or carrier. As used herein "solid phase carrier or support" refers to any support capable of binding the oligonucleotide sequences/DNA regulatory elements of the present invention. Well known supports, or carriers, include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amyloses, natural and modified celluloses, polyacrylamides, agaroses, and magnetite. Methods for coupling nucleic acids to the solid phase, the solid phase substances useful in these methods, and the means for elution of the proteins from the bound ligand, are well known to those of skill in the art.

In addition to the specific methods described above, purification steps prior to affinity separation may also include one or more additional methods, such as ammonium sulfate precipitation, size exclusion chromatography (gel filtration), ion exchange chromatography, differential precipitation and the like, all well known in the art. Also useful is the method known as hydrophobic interaction chromatography (HIC) which is based on the interaction between the solute and the gel that is hydrophobic. Hydrophobic interactions are strongest at high ionic strength, therefore, this form of separation is conveniently performed following salt precipitations or ion exchange procedures. Elution from HIC supports can be effected by alterations in solvent, pH, ionic strength, or by the addition of chaotropic agents or organic modifiers, such as ethylene glycol. General principles of HIC are described in U.S. Patents 3,917,527 and 4,000,098. Purification of specific proteins using HIC is described in the U.S. Patent Nos.: 4,332,717; 4,771, 128; 4,743,680; 4,894,439; 4,908,434; and 4,920,196, the disclosures of which are herein incorporated by reference.

The regulatory elements/oligonucleotide sequences of the present invention may be included in a recombinant DNA construct which contains a regulatory element/oligonucleotide sequence operably linked to a promoter and a heterologous gene. Typically the heterologous gene comprises a reporter sequence, such as the gene for luciferase. In this regard, a recombinant DNA construct, such as a reporter plasmid according to the present invention, can be constructed using conventional molecular biology, microbiology, and recombinant DNA techniques well known to those of skill

WO 95/28482

PCT/US95/04477

24

et al.,

in the art. Such techniques are explained fully in the literature, including Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual" (1982); "DNA Cloning: A Practical Approach," Volumes I and II (D.N. Glover ed. 1985); "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins, eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)]; "Immobilized Cells and Enzymes" [IRL Press, (1986)] and B. Perbal, "A Practical Guide to Molecular Cloning" (1984), the disclosures of which are herein incorporated by reference.

- Promoter sequences useful in DNA constructs according to the present invention include all prokaryotic, eukaryotic or viral promoters capable of driving transcription of a heterologous gene of interest in combination with a regulatory element of the present invention, when transfected into an appropriate host cell.
- Suitable prokaryotic promoters include, but are not limited to, promoters recognized by the T4, T3, Sp6, and T7 polymerases, the P<sub>R</sub> and P<sub>L</sub> promoters of bacteriophage, transcriptional regulatory protein, recA, heat shock, and lacZ promoters of *E. coli*, the -28-specific promoters of *B. subtilis*, the promoters of the bacteriophages of *Bacillus*, *Streptomyces* promoters, the int promoter of bacteriophage, the bla promoter of the  $\beta$ -lactamase gene of pBR322 and the CAT promoter of the chloramphenicol acetyl transferase gene of pPR325. See, e.g., B.R. Glick, 1 *J. Ind. Microbiol.*, 277-282 (1987); Y. Cenatiempo, 68 *Biochimie*, 505-516 (1986); J.D. Watson et al., In: *Molecular Biology of the Gene*, Fourth Edition, Benjamin Cummins, Menlo Park, CA (1987) and S. Gottesman, 18 *Ann. Rev. Genet.*, 415-442 (1984), the disclosures of which are herein incorporated by reference. Preferred eukaryotic promoters include the yeast cyc-1 promoter, the promoter of the mouse metallothionein I gene, the thymidine kinase promoter of the Herpes simplex virus, the SV40 early promoter, and the yeast gal-4 gene promoter. See Guarante et al., 78 *Proc. Natl. Acad. Sci. USA*, 2199-2203 (1981); D. Hamer et al., 1 *J. Mol. Appl. Gen.*, 273-288 (1982); S. McKnight, 31 *Cell*, 355-365 (1982); C. Benoist et al., 290 *Nature* (London), 304-310 (1981); S.A. Johnston et al., 79 *Proc Natl. Acad. Sci. (USA)*, 6971-6975 (1982) and

are useful  
in the  
art

100

WO 95/28482

PCT/US95/04477

25

P.A. Silver et al., 81 Proc. Natl. Acad. Sci. (USA), 5951-5955 (1984), the disclosures of which are herein incorporated by reference herein. Preferably, a DNA construct according to the present invention utilizes the thymidine kinase gene promoter of the Herpes simplex virus.

5 The third component of the recombinant DNA or construct molecules of the present invention is a "heterologous gene" which may be composed of any set of nucleotides regardless of sequence. Nonlimiting examples of such heterologous genes include the structural genes for luciferase,  $\beta$ -galactosidase, chloramphenicol acetyl transferase, secreted placental alkaline phosphatase, human growth hormone, tPA, 10 green fluorescent protein and interferon. For a more extensive list of heterologous genes usable in the constructs and methods of the present invention, see Beaudet, 37 Am. J. Hum. Gen., 386-406 (1985).

15 Preferably the heterologous gene comprises a reporter gene whose product is used to assess regulation of transcription via a promoter and a regulatory element/ oligonucleotide sequence of the present invention. The expression of this "reporter sequence" results in the formation of a reporter product (e.g., protein) which is readily detectable. The reporter sequence will preferably be selected such that the reporter molecule will have a physical and chemical characteristics which will permit or facilitate its identification or detection by means well known in the art. In one 20 embodiment, the presence of the reporter molecule will be detected through the use of an antibody or an antibody fragment, capable of specific binding to the reporter molecule. In another embodiment, a reporter such as  $\beta$ -galactosidase or luciferase can be assayed enzymatically or immunologically.

25 A preferred reporter molecule is LUC, well known in the art. See, e.g., J. R. De-Wet et al., 7 Mol. Cell Bio., 725 (1987). Because this is an insect gene, it is absent from mammalian cells and the enzyme product can be directly assayed in a cell extract. The level of enzyme activity corresponds to the amount of enzyme that was made, which in turn reveals the level of expression. In addition, LUC mRNA may also be measured directly.

WO 95/28482

PCT/US95/04477

26

Typically, a plasmid containing the recombinant DNA molecule of the present invention, including the LUC gene, is introduced into mammalian cells, which are then grown to, at or near confluency. In this regard, any host cell capable of activating one or more transcriptional regulatory proteins in response to an appropriate signaling molecule or molecules can be transfected with the DNA constructs of the present invention. Preferably, such cytokine-responsive host cells comprise mammalian cells, such as HepG2, U937, ME-180, TF-1 and NFS-60 cells.

The reporter cells are treated with a compound or sample suspected of containing a signaling molecule capable of inducing or activating a transcriptional regulatory protein, for example, a sample of growth medium that has been in contact with cells and may contain signaling molecules secreted or released by the cells. The LUC-producing reporter cells are extracted, and the soluble extracts are supplemented with luciferin and ATP. In the presence of these compounds the action of luciferase generates light, which is detected using a luminometer. The amount of light produced is directly related to the amount of luciferase present in the cellular extract.

With a suitable DNA construct transfected into a host cell, the present invention provides a method for the controlled expression of a gene of interest. Thus, application of a signaling molecule, such as cytokine, to transfected host cells can be used to drive the expression of a heterologous gene to yield defined quantities of a desired product, such as human growth hormone, by any of a variety of cell culture and fermentation techniques well known to those of skill in the art.

Alternatively, when the DNA construct comprises a reporter sequence, such as the gene for luciferase, transfection of the DNA construct into a host cell provides a convenient means for measuring the transcriptional activity of a reporter product in response to a signaling molecule, such as a cytokine or a sample of growth medium that has been in contact with cells and may contain signaling molecules secreted or released by the cells.

Importantly, when transcription of LUC is activated by the transcriptional regulatory protein being assayed, LUC synthesis is increased relative to a control lacking the transcriptional regulatory protein. Thus, the amount of LUC

A 110  
anal  
the  
expression  
of the  
reporter  
gene  
determined

WO 95/28482

PCT/US95/04477

27

enzyme produced is an indirect measure of transcription induced by the activated transcriptional regulatory protein binding to the regulatory elements/oligonucleotide sequences of the present invention, which is operably linked to the LUC gene.

When a preferred host cell, such as a HepG2 cell, is transfected with a reporter DNA construct according to the present invention, it can be utilized in assays to detect agonists and antagonists of signaling molecules that induce gene transcription via activated transcriptional regulatory proteins. As used herein, agonists or antagonists of gene transcription include compounds that intervene at any point within the signaling pathway from interaction between the signaling molecule and a cell surface receptor through activation of one or more transcriptional regulatory proteins and binding of the same to DNA regulatory elements, the end result of which is modulation of gene transcription. Further, as used herein, agonists and antagonists of gene transcription also include potentiators of known compounds with such agonist or antagonist properties. Agonists can be detected by contacting the transfected host cell with a compound or mix of compounds and, after a fixed period of time, determining the level of gene expression (e.g., the level of luciferase produced) within the treated cells. This expression level can then be compared to the expression level of the reporter gene in the absence of the compound(s). The difference between the levels of gene expression, if any, indicated whether the compound(s) of interest agonize the activation of intracellular transcriptional regulatory proteins in an analogous fashion to a known agonist signaling molecule, such as a cytokine. Further, the magnitude of the level of reporter product expressed between the treated and untreated cells provides a relative indication of the strength of that compound(s) as an agonist of gene transcription via a transcriptional regulatory protein pathway.

Alternatively, such a transfected host cell can be used to find antagonists of known agonists, e.g., cytokines such as IL-4, of gene transcription utilizing host cells transfected with the DNA constructs according to the present invention. In such an assay, the compound or compounds of interest are contacted with the host cell in conjunction with one or more known agonists (e.g., cytokines) held at a fixed concentration. The extent to which the compound(s) depress the level of gene

A 12

WO 95/28482

PCT/US95/04477

28

expression in the host cell below that available from the host cell in the absence of compounds, but presence of the known agonist, provides an indication and relative strength of the antagonist properties of such compound(s).

- Thus, the present invention provides methods to assay for agonists and antagonists of gene transcription utilizing the regulatory elements/oligonucleotides of the present invention in appropriate DNA constructs and transfected host cells. Further, the agonist and antagonist compounds discovered utilizing these methods can serve as pharmaceutical agents in the intervention of various cytokine-induced disease states and conditions, or to ameliorate disease states caused by cytokine deficiency, such as inflammation, infection, anemia, cytopenia and cancerous or precancerous conditions.

The invention will be further illustrated by reference to the following non-limiting Examples.

#### EXAMPLE 1

##### 15 Reagents

- Oligonucleotides were obtained from Operon Technologies (Alameda, CA). IFN $\gamma$  was the gift of Dr J. E. Darnell (Rockefeller University, New York, NY); commercially available from Genzyme, Cambridge, MA). Recombinant human IL-6, GM-CSF, Oncostatin M, IL-3 and IL-4 were obtained from R&D Systems (Minneapolis, MN). Recombinant IL-13 was obtained from Biosource (Camarillo, CA). Recombinant IL-2 was obtained from Chiron (Emeryville, CA). Recombinant human Epo and G-CSF were from Amgen, Inc. (Thousand Oaks, CA). Protease inhibitors and poly d(I-C) poly d(I-C) were from Boehringer Mannheim (Indianapolis, IN).

25

##### Cells and cell culture

- HepG2 cells were obtained from ATCC (Rockville, MD) and grown in Eagle's Minimum Essential Medium (EMEM, BioWhittaker, Walkersville, MD) supplemented with fetal bovine serum (10% v/v), glutamine (2 mM) and gentamicin sulfate (50  $\mu$ g/mL, BioWhittaker). U937 cells were obtained from Dr. J. E. Darnell

WO 95/28482

PCT/US95/04477

29

- (commercially available from ATCC) and grown in RPMI-1640 (BioWhittaker) supplemented with fetal bovine serum (10% v/v), glutamine (2 mM) and gentamicin sulfate (50 µg/mL). ME-180 cells were obtained from the ATCC and grown in McCoy's 5A (Gibco/BRL, Gaithersburg, MD) supplemented with fetal bovine serum (10% v/v), glutamine (2 mM) and gentamicin sulfate (50 µg/mL). CTLL-2 and TF-1 cells were obtained from the ATCC and grown in RPMI-1640 (BioWhittaker) supplemented with fetal bovine serum (10% v/v), glutamine (2 mM), gentamicin sulfate (50 µg/mL), and IL-2 (CTLL-2, 200 U/mL) or GM-CSF (TF-1, 5 ng/mL). IL-3-dependent NFS-60 cells were obtained from Dr. J. N. Ihle (St. Jude Children's Research Hospital, Memphis, TN) and were maintained in RPMI-1640 supplemented with fetal bovine serum (10% v/v), glutamine (2 mM), gentamicin sulfate (50 µg/mL) and 10% WEHI-3B-conditioned growth medium (to provide IL-3). Factor-independent NFS-60 cells were selected by withdrawing WEHI-3B-conditioned medium from the culture medium. In about two weeks, the cells adjusted to the new growth conditions and proliferated as well as the parental NFS-60 cells. HepG2 and ME-180 cells were treated with cytokines at 50-75% confluency, U937, CTLL-2, TF-1, and NFS-60 cells at a density of  $2 \times 10^5$  -  $1 \times 10^6$ /ml. Cytokines were used at the following concentrations: IFN $\gamma$ , 5 ng/ml, IL-4, 10-30 ng/ml, IL-6, 10 ng/ml, GM-CSF, 5 ng/ml, Epo, 4-6 U/mL, IL-2, 200 U/mL, IL-3, 20 ng/mL, IL-13, 60 ng/mL, and G-CSF, 20 ng/mL.

#### Preparation of Nuclear Extracts and Electrophoretic Mobility Shift Assays

Nuclear extracts were prepared by NP40 lysis as described in H. B. Sadowski and M. Z. Gilman, 362 *Nature*, 79 (1993), the disclosure of which is herein incorporated by reference. Protein concentrations were measured using Bradford dye binding assay (Bio-Rad Laboratories, Hercules, CA). Nuclear extracts were prepared either from untreated HepG2 cells, HepG2 cells treated for 15 min with IFN $\gamma$ , HepG2 cells treated for 15 min with IL-6, untreated U937 cells, U937 cells treated for 30 min with GM-CSF or IL-4, CTLL-2 cells starved of IL-2 for 18 h and then either left untreated or treated for 30 min with IL-2; TF-1 cells starved of GM-CSF for 18 h and

WO 95/28482

PCT/US95/04477

30

then either left untreated or treated for 30 min with Epo, IL-3 or GM-CSF; ME-180 cells either left untreated or treated for 30 min with IL-4 or IL-13; and NFS-60 cells starved of IL-3 for 16-18 h then either left untreated or treated for 10 min with G-CSF, IL-3 or IL-6. The double-stranded probe oligonucleotides used in the Electrophoretic

- 5 Mobility Shift Assays (EMSAs) were formed by annealing oligonucleotides with the sequences:

- |    |                         |                     |
|----|-------------------------|---------------------|
|    | 5'-GATCTGCTTCCGAACGT    | -3' (SEQ ID NO. 54) |
|    | 3'-ACGAAGGCTTGCACTAG    | -5' (SEQ ID NO. 55) |
| 10 | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 56) |
|    | 3'-ACGAAGGGCTTGCACTAG   | -5' (SEQ ID NO. 57) |
|    | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 58) |
| 15 | 3'-ACGAAGGGGCTTGCACTAG  | -5' (SEQ ID NO. 59) |
|    | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 60) |
|    | 3'-ACGAAGGGGGCTTGCACTAG | -5' (SEQ ID NO. 61) |
| 20 | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 62) |
|    | 3'-ACGAAGGGGGCTTGCACTAG | -5' (SEQ ID NO. 63) |
|    | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 64) |
|    | 3'-ACGAAGGGGGCTTGCACTAG | -5' (SEQ ID NO. 65) |
| 25 | 5'-GATCTGCTTCCGGAACGT   | -3' (SEQ ID NO. 66) |
|    | 3'-ACGAAGGCCTTGCACTAG   | -5' (SEQ ID NO. 67) |
|    | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 68) |
| 30 | 3'-ACGAAGGCCTTGCACTAG   | -5' (SEQ ID NO. 69) |
|    | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 70) |
|    | 3'-ACGAAGGGCCTTGCACTAG  | -5' (SEQ ID NO. 71) |
| 35 | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 72) |
|    | 3'-ACGAAGGGGCCTTGCACTAG | -5' (SEQ ID NO. 73) |
|    | 5'-GATCTGCTTCCCGAACGT   | -3' (SEQ ID NO. 74) |
| 40 | 3'-ACGAAGGGGCCTTGCACTAG | -5' (SEQ ID NO. 75) |



WO 95/28482

PCT/US95/04477

31

- 5'-GATCTGCTTCCTGGAACGT -3' (SEQ ID NO. 76)  
 3'- ACGAAGGACCTTGCACTAG -5' (SEQ ID NO. 77)
- 5'-GATCTGCTTCCTTGGAAACGT -3' (SEQ ID NO. 78)  
 5 3'- ACGAAGGAAACCTTGCACTAG -5' (SEQ ID NO. 79)
- 5'-GATCTGCTTCCTTTGGAACGT -3' (SEQ ID NO. 80)  
 3'- ACGAAGGAAACCTTGCACTAG -5' (SEQ ID NO. 81)
- 10 5'-GATCTGCTTCCAGAACGT -3' (SEQ ID NO. 82)  
 3'- ACGAAGGTCTTGCACTAG -5' (SEQ ID NO. 83)
- 5'-GATCTGCTTCCAGAACGT -3' (SEQ ID NO. 84)  
 3'- ACGAAGGGTCTTGCACTAG -5' (SEQ ID NO. 85)
- 15 5'-GATCTGCTTCCCAGAACGT -3' (SEQ ID NO. 86)  
 3'- ACGAAGGGGTCTTGCACTAG -5' (SEQ ID NO. 87)
- 5'-GATCTGCTTCCCCAGAACGT -3' (SEQ ID NO. 88)  
 20 3'- ACGAAGGGGGTCTTGCACTAG -5' (SEQ ID NO. 89)
- 5'-GATCTGCTTCTTGAACGT -3' (SEQ ID NO. 90)  
 3'- ACGAAGAACTTGCACTAG -5' (SEQ ID NO. 91)
- 25 5'-GATCTGCTTCTTTGAACGT -3' (SEQ ID NO. 92)  
 3'- ACGAAGAAACTTGCACTAG -5' (SEQ ID NO. 93)
- 5'-GATCTGCTTCTTTGAACGT -3' (SEQ ID NO. 94)  
 3'- ACGAAGAAAACTTGCACTAG -5' (SEQ ID NO. 95)
- 30 5'-GATCTGCTTCTTTTGAACGT -3' (SEQ ID NO. 96)  
 3'- ACGAAGAAAAACTTGCACTAG -5' (SEQ ID NO. 97)
- 5'-GATCTGCTTCTAGAACGT -3' (SEQ ID NO. 98)  
 35 3'- ACGAAGATCTTGCACTAG -5' (SEQ ID NO. 99)
- 5'-GATCTGCTTCTCAGAACGT -3' (SEQ ID NO. 100)  
 3'- ACGAAGAGTCTTGCACTAG -5' (SEQ ID NO. 101)
- 40 5'-GATCTGCTTCTCCAGAACGT -3' (SEQ ID NO. 102)  
 3'- ACGAAGAGGTCTTGCACTAG -5' (SEQ ID NO. 103)
- 5'-GATCTGCTTCTCCAGAACGT -3' (SEQ ID NO. 104)  
 3'- ACGAAGAGGGTCTTGCACTAG -5' (SEQ ID NO. 105)

WO 95/28482

PCT/US95/04477

32

5'-GATCGATTTCCTCCGAAATG -3' (SEQ ID NO. 106)  
 3'-CTAAAGGGGCTTTACCTAG -5' (SEQ ID NO. 107)  
 5'-GATCCAATTTCTAAGAAAGGA -3' (SEQ ID NO. 108)  
 3'-GTAAAGATTCTTCTCTAG -5' (SEQ ID NO. 109)  
 5'-GATCCACTTCCCAAGAACAGA -3' (SEQ ID NO. 110)  
 3'-GTGAAGGGTCTTGTCTCTAG -5' (SEQ ID NO. 111)

where the nucleotide sequences shown in bold type face correspond to nucleotide sequences, including their double-stranded complement, tested for activity as regulatory elements according to the present invention.

The annealed oligonucleotides were labeled by filling in the overhanging ends with Klenow fragment (Boehringer Mannheim) in the presence of [ $\alpha$ - $^{32}$ P]-dGTP and/or [ $\alpha$ - $^{32}$ P]-dATP (Amersham Corporation, Arlington Heights, IL). Electrophoretic mobility shift assays (EMSA's) were performed in HEPES buffer (13 mM, pH 7.6, Sigma Chemical, St. Louis, MO), containing sodium chloride (80 mM), sodium fluoride (3mM), sodium molybdate (3 mM), DTT (1mM), EDTA (0.15mM), EGTA (0.15mM), glycerol (8% v/v, including contributions from the nuclear extract), poly d(I-C) poly d(I-C) (75 $\mu$ g/mL), radiolabeled probe (approximately 0.2ng) and nuclear extract containing 5-10 $\mu$ g of total protein. Reactions were incubated at room temperature for 20 minutes then resolved on 5% polyacrylamide gels containing 0.25X TBE [1X TBE is Tris borate (89 mM), pH 8.0 containing EDTA (1 mM)] and glycerol (5% v/v). Gels were run at 4°C in 0.25X TBE at 20V/cm, then dried and autoradiographed.

Relative binding affinities, as determined from the EMSA results for oligonucleotide SEQ ID NOs 54-105, were visually rated and assigned according to the following scale:

30

WO 95/28482

PCT/US95/04477

33

- (-) band corresponding to specific complex on the EMSA autoradiogram (See e.g., FIG. 2C, lanes 2 and 3 ) barely discernible or not discernible.
- (+) band corresponding to specific complex on the EMSA autoradiogram (See e.g., FIG. 2A, lane 3) easily discernible but of weak intensity.
- 5 (++) band corresponding to specific complex on the EMSA autoradiogram (See e.g., FIG. 2B, lane 5) easily discernible and of moderate intensity.
- (+++ band corresponding to specific complex on the EMSA autoradiogram (See e.g., FIG. 2C, lane 6) easily discernible and of strong intensity.

- 10 This visual rating system is sufficient to analyze distinguishable differences and trends in the EMSA binding data as opposed to specific numerical values. If desired, the use of a phosphor imager or densitometer (commercially available from e.g., Bio-Rad Laboratories) could provide a means to assess the differences described here quantitatively. Specific visual ratings of binding affinities
- 15 for the regulatory elements of oligonucleotide SEQ ID NOs 54-105 are shown in Tables 1-6 below (in the Tables, bold type face highlight the spacing nucleotides).

- 20 [rest of page left blank on purpose]

WO 95/28482

PCT/US95/04477

34

Table 1: Relative EMSA binding affinities for a series of regulatory elements in double stranded configurations (SEQ ID NOs. 54-65) having between 3-8 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing (x) | Core Regulatory Element Series<br>TTN <sub>x</sub> AA | IFN $\gamma$ | IL-6 | GM  | IL-4 |
|-------------|---|--------------|------|-----|------|
| 3           | TTCCGAA   | -            | -    | -   | -    |
| 4           | TTCCCGAA  | -            | +++  | -   | -    |
| 5           | TTCCCCGAA   | +            | -    | +++ | ++   |
| 6           | TTCCCCCGAA  | -            | -    | -   | +++  |
| 7           | TTCCCCCGGAA   | -            | -    | -   | -    |
| 8           | TTCCCCCCGAA   | -            | -    | -   | -    |

Table 2: Relative EMSA binding affinities for a series of regulatory elements in double stranded configurations (SEQ ID NOs. 66-75) having between 4-8 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing (x) | Core Regulatory Element Series<br>TTN <sub>x</sub> AA | IFN $\gamma$ | IL-6 | GM  | IL-4 |
|-------------|---|--------------|------|-----|------|
| 4           | TTCCGGAA  | -            | +    | -   | -    |
| 5           | TTCCCGGAA   | ++           | +++  | +++ | ++   |
| 6           | TTCCCCGGAA  | -            | -    | -   | +++  |
| 7           | TTCCCCCGGAA   | -            | -    | -   | -    |
| 8           | TTCCCCCCGGAA  | -            | -    | -   | -    |

WO 95/28482

PCT/US95/04477

35

Table 3: Relative EMSA binding affinities for a series of regulatory elements in double stranded configurations (SEQ ID NOs. 66-67 and 76-81) having between 4-7 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing<br>(x) | Core Regulatory<br>Element Series<br><u>TTN<sub>x</sub>AA</u> | IFN $\gamma$ | IL-6 | GM | IL-4 |
|----------------|---|--------------|------|----|------|
| 4              | TTCCGGAA  | -            | +    | -  | -    |
| 5              | TTCTTGAA  | +            | ++   | ++ | ++   |
| 6              | TTCTTGAA  | -            | -    | -  | +++  |
| 7              | TTCTTTGAA   | -            | -    | -  | -    |

Table 4: Relative EMSA binding affinities for a series of regulatory elements in double stranded configurations (SEQ ID NOs. 82-89) having between 4-7 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing<br>(x) | Core Regulatory<br>Element Series<br><u>TTN<sub>x</sub>AA</u> | IFN $\gamma$ | IL-6 | GM  | IL-4 |
|----------------|---|--------------|------|-----|------|
| 4              | TTCCAGAA  | -            | +    | -   | -    |
| 5              | TTCCCAGAA   | +            | ++   | +++ | ++   |
| 6              | TTCCCCAGAA  | -            | -    | -   | +++  |
| 7              | TTCCCCAGAA  | -            | -    | -   | -    |

WO 95/28482

PCT/US95/04477

36

Table 5: Relative EMSA binding affinities for a series of regulatory elements in double stranded configurations (SEQ ID NOs. 90-97) having between 4-7 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing (x) | Core Regulatory Element Series<br>TTN <sub>x</sub> AA | IFN $\gamma$ | IL-6 | GM | IL-4 |
|-------------|---|--------------|------|----|------|
| 4           | TTCTTGAA  | -            | -    | -  | -    |
| 5           | TTCTTTGAA   | -            | -    | +  | +    |
| 6           | TTCTTTTGAA  | -            | -    | -  | +    |
| 7           | TTCTTTTGAA  | -            | -    | -  | -    |

Table 6: Relative EMSA binding affinities for a series of regulatory elements in double stranded configurations (SEQ ID NOs. 98-105) having between 4-7 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing (x) | Core Regulatory Element Series<br>TTN <sub>x</sub> AA | IFN $\gamma$ | IL-6 | GM | IL-4 |
|-------------|---|--------------|------|----|------|
| 4           | TTCTAGAA  | -            | -    | -  | -    |
| 5           | TTCTCAGAA   | -            | -    | ++ | ++   |
| 6           | TTCTCCAGAA  | -            | -    | -  | ++   |
| 7           | TTCTCCAGAA  | -            | -    | -  | -    |

In evaluating the data in Tables 1-6, one should be aware that IL-6

- induces three specific DNA-binding complexes in HepG2 cells. P. Lamb et al., 83 *Blood* 2063 (1994). Accordingly, for the columns marked "IL-6," the complex being rated is the slowest-migrating IL-6-induced complex in the EMSA, i.e., the STAT3

homodimer (Z. Zhong et al., 264 *Science* 95), and thus is not reflective of complexes that contain STAT1 $\alpha$ , which is rated for the IFN $\gamma$  complex. Further, when interpreting the results contained in Tables 1-6, the data should be analyzed for trends as opposed to specific numerical values. This is due to the inherent lack of fine sensitivity in the EMSA assay. This variability arises, at least in part, from differences in the quality of the nuclear extracts used, cell line differences, and variability in the protein concentration determinations.

The binding data summarized in Tables 1-6 show that DNA regulatory element sequences comprising different nucleotide spacing between the palindromic element (TT-space-AA) have surprisingly different affinities for the activated transcriptional regulatory proteins activated by IFN $\gamma$ , IL-6, GM-CSF and IL-4 treatment of responsive host cells. Specifically, the four base pair spacing shows remarkable selectivity in its binding to the activated transcriptional regulatory protein complex induced by IL-6, i.e., STAT3. Further, this will hold true for other signaling molecules that activate STAT3, such as Oncostatin M, LIF, EGF, PDGF, G-CSF, IL-10, IL-11 and CNTF. The five base pair spacing shows little selectivity between the activated transcriptional regulatory protein complexes, but is noteworthy due to its affinity in most cases for all of the cytokine-induced complexes tested. The six base pair spacing again shows remarkable selectivity, but, in this case, in its binding to the activated transcriptional regulatory protein complex(es) induced by IL-4 (i.e. STAT5 and/or STAT6). The binding data further show that the exact identity of the spacing nucleotides also has some influence on binding affinity, i.e. spacing nucleotides rich in thymines and deoxyadenosines (SEQ ID NOs. 90-105) disfavor STAT1 $\alpha$  and STAT3 binding (Tables 4-6). Thus, the elements with a five base pair spacer rich in thymines and deoxyadenosines were selective for the STAT complexes activated by GM-CSF and IL-4 (i.e., STAT 5 and/or STAT6) over those containing STAT1 $\alpha$  or STAT3.

For those oligonucleotides that were tested, the spacing preference of the STAT complexes activated by IL-2 in CTLL-2 cells, Epo in TF-1 cells, and IL-3 in NFS-60 cells is the same as for the complexes activated by GM-CSF in U937 cells.

This is in accord with the fact that IL-2, GM-CSF, Epo and IL-3 all are reported to activate the same STAT protein, STAT5. The spacing preference for the STAT complexes activated by IL-13 in ME-180 cells was the same as for the STAT complexes activated by IL-4 in U-937 and ME-180 cells, and is due to the likelihood that IL-4 and IL-13 both activate the same STAT protein, STAT6.

In NFS-60 cells, G-CSF activated two STAT complexes that were distinguishable by their differing mobilities in an EMSA. The slower-migrating complex comigrated with the STAT3 homodimer stimulated by IL-6 and was shown to contain STAT3 by antibody supershift experiments using a specific STAT3 antiserum. The faster-migrating complex contained an unidentified STAT complex that migrated like the STAT complexes activated by IL-3 and GM-CSF. In binding experiments, for the oligonucleotide series that were tested, the G-CSF-activated STAT3-like complex had spacing preferences indistinguishable from the STAT3 homodimer induced by IL-6. Similarly, the G-CSF-activated STAT complex that comigrated with the GM-CSF- and IL-3-activated STAT complexes had the same corresponding spacing preferences.

Data are not shown in Tables 1-6 for the regulatory elements SEQ ID NOs. 106, 108 and 110. SEQ ID NO. 106 is a regulatory element that has a five base pair spacing and will bind to the STAT complexes activated by IFN $\gamma$  (+++), IL-6 (+++), GM-CSF (+++) and IL-4 (+++). SEQ ID NO. 108 is a regulatory element that has a five base pair spacing and will bind to the STAT complexes activated by IFN $\gamma$  (+++), IL-6 (+), GM-CSF (+++) and IL-4 (+). SEQ ID NO. 110 is a regulatory element that has a six base pair spacing and selectively bound to the STAT complex activated by IL-4 (+++), but not to STAT complexes activated by IFN $\gamma$ , IL-6 or GM-CSF. Thus, when tested for *in vitro* binding, SEQ ID NOs. 106, 108 and 110 also fit the spacing patterns described above.

A specific example of this nucleotide spacing effect can be seen with respect to the EMSA autoradiograms of FIGS. 2A-2D. In panel 2A, a radiolabeled, double-stranded oligonucleotide probe (corresponding to a four base pair spacer) made by annealing oligonucleotides of SEQ ID NOs. 66 and 67 was used. Lanes marked 'N'



WO 95/28482

PCT/US95/04477

39

represent experiments using extracts from untreated cells. Other lanes are marked according to the inducing cytokine. Activated complexes can be identified by their absence in untreated extracts and their presence in extracts treated by cytokine. The only activated transcriptional regulatory protein complex that binds to this four-spacer probe can be seen in lane 3 (identified on FIG. 2A as the STAT3 homodimer), and corresponding to the IL-6-induced extract.

In panel 2B, a radiolabeled, double-stranded oligonucleotide probe (corresponding to a five base pair spacer) made by annealing oligonucleotides of SEQ ID NOs 76 and 77 was used. Lanes marked 'UN' represent experiments using extracts from untreated cells. Other lanes are marked according to the inducing cytokine. This panel shows that IFN $\gamma$ , IL-6, GM-CSF and IL-4 all activate transcriptional regulatory protein complexes that are bound by this probe. These activated complexes can be seen in lanes 2, 3, 5 and 6 and the various complexes are identified on FIG. 2B.

In panel 2C, a radiolabeled, double-stranded oligonucleotide probe (corresponding to a six base pair spacer) made by annealing oligonucleotides of SEQ ID NOs 78 and 79 was used. Lanes marked 'UN' represent experiments using extracts from untreated cells. Other lanes are marked according to the inducing cytokine. The only activated transcriptional regulatory protein complex that binds to this six-spacer probe can be seen in lane 6 (identified FIG. 2C), the IL-4-induced extract.

In panel 2D, a radiolabeled, double-stranded oligonucleotide probe (corresponding to a seven base pair spacer) made by annealing oligonucleotides of SEQ ID NOs 80 and 81 was used. Lanes marked 'UN' represent experiments using extracts from untreated cells. Other lanes are marked according to the inducing cytokine. Only complexes that bind constitutively (marked on FIG. 2D) can be detected using this seven-spacer probe; no cytokine-activated transcriptional regulatory protein complexes are apparent.

#### Transient transfection assays

The reporter plasmids SEQID56x4TK-LUC, SEQID68x4TK-LUC, SEQID106x4TK-LUC, SEQID70x4TK-LUC, SEQID108x6TK-LUC, and

SEQID110x4TK-LUC contain four copies (or six copies for SEQID108x6TK-LUC) of oligonucleotide sequences with the indicated SEQ ID NOs linked to the promoter of the Herpes Simplex virus thymidine kinase gene at position -35 with respect to the cap site. See FIG. 1. The reference reporter, TK-LUC (P. Lamb et al., 8 Blood 2063 (1994)), the disclosure of which is herein incorporated by reference, is the parent vector that contains no regulatory element. These chimeric promoters drive the expression of the structural gene for firefly luciferase.

HepG2 and ME-180 cells were transfected with the reporter plasmids of above by calcium phosphate coprecipitation. Cells were seeded at  $1-4 \times 10^5$ /ml the day before transfection. Cells were exposed to a calcium phosphate precipitate containing the above reporter plasmids (10-20  $\mu$ g/ml) and the  $\beta$ -galactosidase-expressing plasmid pCHI110 (5  $\mu$ g/ml, commercially available from Pharmacia Biotech, Piscataway, NJ) for 6 h (HepG2) or 12 h (ME-180). The medium was then changed and the cells allowed to recover for 16-18 h. Recombinant cytokines were then added prediluted in growth medium and the cells harvested after 5 h (HepG2) or 6 h (ME-180). Cells were lysed and luciferase and  $\beta$ -galactosidase activities determined using standard techniques. See, e.g. J. R. De Wet et al., 7 Mol. Cell. Biol. 725 (1987) and Sambrook et al., Molecular Cloning: A Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), the disclosure of which is herein incorporated by reference. For each sample the normalized response was determined by dividing light units obtained from the luciferase assay with the  $\beta$ -galactosidase activity in the same lysate as determined using a chromogenic substrate. The results of these transfections are shown below in Tables 7-8. Numbers given are the mean fold inductions ('fold induction' is defined as the normalized response in a cytokine-treated sample divided by the normalized response in an untreated sample).

TF-1 cells were transfected with the reporter plasmids of above by the DEAE-dextran method as described by J. Suzow and A. D. Friedman, 13 Mol. Cell. Biol., 2141 (1993), the disclosure of which is herein incorporated by reference, with the following modifications: test reporter constructs were added to a concentration of 3

WO 95/28482

PCT/US95/04477

41

µg/mL during the transfection, pMSVCAT vector was not added to the transfection mixtures, the growth medium used was as described above for TF-1 cells, and cytokine inductions were carried out for 4-5 h. Cells were lysed and luciferase activity determined using standard techniques. Transfections were performed in a batch, and identical numbers of transfected cells were then separately induced with cytokine for the 4-5 h induction period. Numbers given are the mean fold inductions ('fold induction' for TF-1 transfections is defined as the luciferase response in a cytokine-treated sample divided by the luciferase response in an untreated sample). The results of these transfections are shown in Table 9.

NFS-60 cells were transfected with the reporter plasmids described above by the DEAE-dextran method as described in the preceding paragraph with the following modifications: test reporter constructs were added to a concentration of 6 µg/mL during the transfection, and cytokine inductions were carried out for 2.5h. The results of these transfections are shown in Table 10.

Table 7: Transcriptional Induction in HepG2 Cells of Reporter Constructs Incorporating Multiple Copies of STAT-Selective Regulatory Elements. The values given are mean fold inductions in response to the indicated cytokine. The value in the parentheses is the number of experiments included to calculate the mean.

| Reporter         | Core Element | IFN $\gamma$ | IL-6    | Oncostatin M |
|------------------|--------------|--------------|---------|--------------|
| TK-LUC           | none         | 1.0 (3)      | 1.8 (3) | 2.9 (3)      |
| SEQID56x4TK-LUC  | TTCCCGAA     | 0.8 (3)      | 9.7 (3) | 20 (3)       |
| SEQID68x2TK-LUC  | TTCCCGAA     | 34 (3)       | 19 (3)  | 30 (3)       |
| SEQID106x4TK-LUC | TTCCCGAA     | 60 (3)       | 43 (3)  | 90 (3)       |
| SEQID70x4TK-LUC  | TTCCCGGAA    | 0.8 (3)      | 0.9 (3) | 1.0 (3)      |
| SEQID110x4TK-LUC | TTCCCAAGAA   | 0.9 (3)      | 1.8 (3) | 3.2 (3)      |

WO 95/28482

PCT/US95/04477

42

Table 8: Transcriptional Induction in ME-180 Cells of Reporter Constructs Incorporating Multiple Copies of STAT-Selective Regulatory Elements. The values given are mean fold inductions in response to the indicated cytokine. The value in the parentheses is the number of experiments included to calculate the mean.

| Reporter         | Core Element | IFN $\gamma$ | IL-6    | Oncostatin M | IL-4    |
|------------------|--------------|--------------|---------|--------------|---------|
| TK-LUC           | none         | 1.0 (3)      | 1.2 (3) | 1.3 (3)      | 0.9 (3) |
| SEQID56x4TK-LUC  | TTCCCGAA     | 1.2 (3)      | 2.4 (3) | 4.3 (3)      | 1.4 (3) |
| SEQID68x2TK-LUC  | TTCCCGAA     | 3.7 (3)      | 3.0 (3) | 5.1 (3)      | 1.2 (3) |
| SEQID106x4TK-LUC | TTCCCGAA     | 5.5 (3)      | 4.5 (3) | 11 (3)       | 1.0 (3) |
| SEQID70x4TK-LUC  | TTCCCGAA     | 1.1 (3)      | 1.3 (3) | 1.4 (3)      | 1.2 (3) |
| SEQID110x4TK-LUC | TTCCAAGAA    | 1.0 (3)      | 1.4 (3) | 1.2 (3)      | 8.1 (3) |

Table 9: Transcriptional Induction in TF-1 Cells of Reporter Constructs Incorporating Multiple Copies of STAT-Selective Regulatory Elements. The values given are mean fold inductions in response to the indicated cytokine. The value in the parentheses is the number of experiments included to calculate the mean.

| Reporter         | Core Element | IL-4    | IL-6    | GM-CSF  | Epo     | IL-3    |
|------------------|--------------|---------|---------|---------|---------|---------|
| TK-LUC           | none         | 0.7 (2) | 1.2 (2) | 0.8 (2) | 0.8 (2) | 0.7 (2) |
| SEQID56x4TK-LUC  | TTCCCGAA     | 0.7 (1) | 5.5 (2) | 1.2 (2) | 0.6 (2) | n.d.    |
| SEQID68x4TK-LUC  | TTCCCGAA     | n.d.    | 19 (1)  | 3.4 (2) | 1.8 (2) | 2.8 (2) |
| SEQID106x4TK-LUC | TTCCCGAA     | 1.4 (2) | 43 (3)  | 9.6 (4) | 3.1 (4) | 6.8 (3) |
| SEQID108x6TK-LUC | TTCTAAGAA    | n.d.    | 3.1 (1) | 7.6 (2) | 3.5 (2) | 7.3 (2) |
| SEQID110x4TK-LUC | TTCCAAGAA    | 3.3 (3) | 1.3 (1) | 1.3 (2) | 0.8 (2) | 1.3 (2) |

n.d. = not determined

WO 95/28482

PCT/US95/04477

43

Table 10: Transcriptional Induction in NFS-60 Cells of Reporter Constructs Incorporating Multiple Copies of STAT-Selective Regulatory Elements. The values given are mean fold inductions in response to the indicated cytokine. The value in the parentheses is the number of experiments included to calculate the mean.

| Reporter             | Core Element | G-CSF    | IL-3    | IL-6    |
|----------------------|--------------|----------|---------|---------|
| SEQID56x4<br>TK-LUC  | TTCCCGAA     | 13.8 (2) | 1.0 (2) | 4.1 (2) |
| SEQID68x2<br>TK-LUC  | TTCCCGGAA    | 23.8 (2) | 3.5 (2) | 4.8 (2) |
| SEQID106x4<br>TK-LUC | TTCCCGGAA    | 23.6 (2) | 6.1 (2) | 6.9 (2) |
| SEQID110x4<br>TK-LUC | TTCCCAAGAA   | 1.0 (1)  | 1.0 (1) | n.d.    |

n.d. = not determined

One of the regulatory elements that was incorporated into the TK-LUC reporter construct used to generate the data in Tables 7-10 (SEQID 106) was taken from the promoter of the human IRF-1 gene, a gene reported to be regulated by IFN $\gamma$ , and, in light of the teaching of the present invention, can be seen as a natural counterpart to the five base pair synthetic spacing elements in Tables 1-6. One of the regulatory elements that was incorporated into the TK-LUC reporter construct used to generate the data in Tables 7-9 (SEQID 108) was taken from the promoter of the human Fc $\epsilon$ RIIb gene, and, in light of the teaching of the present invention, can be seen as a natural counterpart to the five base pair synthetic spacing elements in Tables 1-6. One of the regulatory elements that was incorporated into the TK-LUC reporter construct used to generate the data in Tables 7-10 (SEQID 110) was taken from the promoter of the human C $\epsilon$  gene, a gene reported to be regulated by IL-4, and, in light of the teaching of the present invention, can be seen as a natural counterpart to the six base pair synthetic spacing elements in Tables 1-6.

Table 7 shows that the binding data for the IFN $\gamma$ - and IL-6-activated transcriptional regulatory protein complexes shown in Tables 1-6 correlate with selective activation of transcription. Thus, a four nucleotide spacing regulatory element

not only selectively binds the IL-6-induced activated transcriptional regulatory protein complex but, when placed in a functional promoter upstream of a reporter gene, selectively activates the reporter gene transcription. Oncostatin M also activates STAT3 and thus also stimulates transcription from the palindromic element with a spacing of four (SEQ ID NO. 56), as will other signaling molecules that activate STAT3. The five nucleotide spacing elements (SEQ ID NOs. 68 and 106) bind both the IFN $\gamma$ - and IL-6-induced transcriptional regulatory protein complexes and, correspondingly, also confer responsiveness to IFN $\gamma$ , IL-6 and Oncostatin M when placed in the requisite reporter construct. The six nucleotide spacing elements (SEQ ID NOs. 70 and 110) bind the IFN $\gamma$ - and IL-6-induced activated transcriptional regulatory protein complexes weakly, and, correspondingly do not confer responsiveness in the context of a reporter construct to IFN $\gamma$ , and confers much lower responsiveness in the context of a reporter construct to IL-6 and Oncostatin M. Thus, the binding selectivity of the palindromic elements correlates directly with the transcriptional activation potential of the regulatory elements of the present invention.

The transfection results for the reporter constructs in ME-180 cells are shown in Table 8. The element with a four base pair spacing (SEQ ID NO. 56) again selectively mediated transcriptional induction by IL-6 and OSM, activators of STAT3. The reporters containing the five base pair spacing elements (SEQ ID NOs. 68, 106 and 108) again responded to IFN $\gamma$ , IL-6, and OSM; however, they did not respond to IL-4 though they bound to the IL-4-activated STAT complex. The reporter containing the six base pair synthetic spacing element (SEQ ID NO. 70) did not respond to any of the cytokines tested (though it bound specifically to the IL-4-activated STATs), but the natural C $\epsilon$  element with a six base pair spacing (SEQ ID NO. 110) mediated strong and extremely selective transcriptional induction by IL-4, and is therefore a functional activation site in accord with the *in vitro* binding data.

The transfection results for reporter constructs in TF-1 cells are shown in Table 9. The results are substantially consistent with the *in vitro* binding data described above. Thus, the element with a four base pair spacing (SEQ ID NO. 56) was

WO 95/28482

PCT/US95/04477

45

unresponsive to IL-4, Epo, IL-3 and GM-CSF but did respond to IL-6, while the reporters containing the five base pair spacing elements (SEQ ID NOs. 68, 106 and 108) responded to IL-6, Epo, IL-3 and GM-CSF but not to IL-4. Accordingly, the reporter containing the natural six base pair spacing element (SEQ ID NO. 110) responded to IL-4 but not to IL-6, Epo, IL-3 or GM-CSF.

The transfection results for the reporter constructs tested in NFS-60 cells are shown in Table 10. The results are substantially consistent with the *in vitro* binding data described above. Thus, the element with a four base pair spacing (SEQ ID NO. 56) mediated a response to both IL-6 and G-CSF (both of which activate STAT3), while the elements with a five base pair spacing (SEQ ID NOs. 68 and 108), both of which bind to both STAT complexes activated by G-CSF, mediated a response to IL-3, IL-6 and G-CSF (in accord with the *in vitro* binding data). Accordingly, the reporter containing the natural six base pair spacing element (SEQ ID NO. 110) did not respond to IL-3 or G-CSF.

## EXAMPLE 2

Preparation of the reagents, cells, cell culture, nuclear extracts and Electrophoretic Mobility Shift Assays were performed as described in Example 1, with the exception that the double-stranded probe oligonucleotides used in the Electrophoretic Mobility Shift Assays (EMSAs) were formed by annealing oligonucleotides with the sequences:

|                         |                      |
|-------------------------|----------------------|
| 5'-GATCTGCTTACGTAACGT   | -3' (SEQ ID NO. 112) |
| 3'-ACGAATGCATTGCACTAG   | -5' (SEQ ID NO. 113) |
| 5'-GATCTGCTTACCGTAACGT  | -3' (SEQ ID NO. 114) |
| 3'-ACGAATGGCATTGCACTAG  | -5' (SEQ ID NO. 115) |
| 5'-GATCTGCTTACCCGTAACGT | -3' (SEQ ID NO. 116) |
| 3'-ACGAATGGGCATTGCACTAG | -5' (SEQ ID NO. 117) |
| 5'-GATCTGCTTACCCGTAACGT | -3' (SEQ ID NO. 118) |
| 3'-ACGAATGGGCATTGCACTAG | -5' (SEQ ID NO. 119) |

WO 95/28482

PCT/US95/04477

46

|    |                          |                      |
|----|--------------------------|----------------------|
|    | 5'-GATCTGCTTCCGTAACGT    | -3' (SEQ ID NO. 120) |
|    | 3'-ACGAAGGCATTGCACTAG    | -5' (SEQ ID NO. 121) |
|    | 5'-GATCTGCTTCCCGTAACGT   | -3' (SEQ ID NO. 122) |
| 5  | 3'-ACGAAGGGCATTGCACTAG   | -5' (SEQ ID NO. 123) |
|    | 5'-GATCTGCTTCCCCGTAACGT  | -3' (SEQ ID NO. 124) |
|    | 3'-ACGAAGGGGCATTGCACTAG  | -5' (SEQ ID NO. 125) |
| 10 | 5'-GATCTGCTTCCCCGTAACGT  | -3' (SEQ ID NO. 126) |
|    | 3'-ACGAAGGGGCATTGCACTAG  | -5' (SEQ ID NO. 127) |
|    | 5'-GATCTGCTTCTGTAACGT    | -3' (SEQ ID NO. 128) |
| 15 | 3'-ACGAAGACATTGCACTAG    | -5' (SEQ ID NO. 129) |
|    | 5'-GATCTGCTTCTCGTAACGT   | -3' (SEQ ID NO. 130) |
|    | 3'-ACGAAGAGCATTGCACTAG   | -5' (SEQ ID NO. 131) |
|    | 5'-GATCTGCTTCTCCGTAACGT  | -3' (SEQ ID NO. 132) |
| 20 | 3'-ACGAAGAGGCATTGCACTAG  | -5' (SEQ ID NO. 133) |
|    | 5'-GATCTGCTTCTCCGTAACGT  | -3' (SEQ ID NO. 134) |
|    | 3'-ACGAAGAGGGCATTGCACTAG | -5' (SEQ ID NO. 135) |

25 where the nucleotide sequences shown in bold type face correspond to nucleotide sequences, including their double-stranded complement, tested for activity as regulatory elements according to the present invention.

Relative binding affinities, as determined from the EMSA results for oligonucleotide SEQ ID NOs 112-135, were visually rated and assigned according to the same visual rating scale as described in Example 1.

30 This visual rating system is sufficient to analyze distinguishable differences and trends in the EMSA binding data as opposed to specific numerical values. If desired, the use of a phosphor imager or densitometer (commercially available from e.g., Bio-Rad Laboratories) could provide a means to assess the differences described here quantitatively. Specific visual ratings of binding affinities for the regulatory elements of oligonucleotide SEQ ID NOs 112-135 are shown in Tables 11-13 below (in the Tables, bold type face highlight the spacing nucleotides).



WO 95/28482

PCT/US95/04477

47

Table 11: Relative EMSA binding affinities for a series of regulatory elements of double stranded configurations (SEQ ID NOs. 114-121) having between 4-7 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing (x) | Core Regulatory Element Series<br>TTN <sub>x</sub> AA | IFN $\gamma$ | IL-6 | GM | IL-4 |
|-------------|---|--------------|------|----|------|
| 4           | TTACGTAA  | -            | -    | -  | -    |
| 5           | TTACCGTAA   | +            | ++   | -  | -    |
| 6           | TTACCCGTAA  | -            | -    | +  | +    |
| 7           | TTACCCCGTAA   | -            | -    | -  | ++   |

Table 12: Relative EMSA binding affinities for a series of regulatory elements of double stranded configurations (SEQ ID NOs. 122-129) having between 4-7 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing (x) | Core Regulatory Element Series<br>TTN <sub>x</sub> AA | IFN $\gamma$ | IL-6 | GM | IL-4 |
|-------------|---|--------------|------|----|------|
| 4           | TTCCGTAA  | -            | -    | -  | -    |
| 5           | TTCCCGTAA   | +            | ++   | -  | -    |
| 6           | TTCCCCGTAA  | -            | -    | +  | +    |
| 7           | TTCCCCCGTAA   | -            | -    | -  | +    |

Table 13: Relative EMSA binding affinities for a series of regulatory elements of double stranded configurations (SEQ ID NOs. 130-137) having between 4-7 spacing deoxynucleotides in response to the cytokines IFN $\gamma$ , IL-6, GM-CSF (GM=GM-CSF) and IL-4.

| Spacing (x) | Core Regulatory Element Series<br>TTN <sub>x</sub> AA | IFN $\gamma$ | IL-6 | GM | IL-4 |
|-------------|---|--------------|------|----|------|
| 4           | TTCTGTAA  | -            | +    | -  | -    |
| 5           | TTCTCGTAA   | +            | ++   | -  | -    |
| 6           | TTCTCCGTAA  | -            | -    | ++ | +    |
| 7           | TTCTCCCGTAA   | -            | -    | -  | ++   |

In evaluating the data in Tables 11-13, one should be aware of the that IL-6 induces three specific DNA-binding complexes in HepG2 cells. P. Lamb et al., 8 *Blood* 2063 (1994). Accordingly, for the columns marked "IL-6," the complex being rated is the slowest-migrating IL-6-induced complex in the EMSA, i.e., the STAT3 homodimer (Z. Zhong et al., 264 *Science* 95), and thus is not reflective of complexes that contain STAT1, which is rated for the IFN $\gamma$  complex. Further, when interpreting the results contained in Tables 11-13, the data should be analyzed for trends as opposed to specific numerical values. This is due to the inherent lack of fine sensitivity in the EMSA assay. This variability arises, at least in part, from differences in the quality of the nuclear extracts used, cell line differences, and variability in the protein concentration determinations. The binding affinities of STAT complexes for the oligonucleotides in Tables 11-13 were generally lower (by roughly 10-fold or more) than for the series described in Example 1.

The binding data summarized in Tables 11-13 show that DNA regulatory element sequences comprising different nucleotide spacings between the palindromic (TT-space-AA) element have surprisingly different affinities for the activated transcriptional regulatory proteins activated by IFN $\gamma$ , IL-6, GM-CSF and IL-4 treatment

of responsive cells. For the spacing elements where one or both core half sites is 'TTA', the spacing preferences of some STAT complexes are subtly different from those described in Example 1. In general, the four base pair spacing did not bind any complexes except in one instance (Table 13) in which the STAT3-containing complex (IL-6-activated) was selectively bound, consistent with the data shown in Example 1. STAT1- and STAT3-containing complexes (activated by IFN $\gamma$  or IL-6) bound to the five base pair spacing in each series. For these series of 'TTA' core half site-containing elements, the binding preference for the STAT complexes activated by GM-CSF is a six base pair spacing (instead of five base pair spacing as shown in Example 1). The IL-4-activated STAT complexes also can bind to these six base pair spacing elements, but now the seven base pair spacing selectively binds the IL-4-activated STAT complexes. Therefore, unlike for STAT1- and STAT3-containing complexes, the spacing preference for the GM-CSF- and IL-4-activated STAT complexes has been expanded by one nucleotide when one or both of the core half sites is 5'-TTA-3'. This shift in spacing preference rendered the regulatory-elements in this Example 2 with a spacing of five base pair selective for STAT1 and/or STAT3 over the STAT complexes activated by GM-CSF or IL-4. Thus, the nature of the internal nucleotides can affect not only the absolute binding affinity of a given STAT complex (c.f. Tables 4-6 and 11-13) but also can affect the spacing preference of the various STAT complexes (Tables 11-13).

For those oligonucleotides that were tested, the spacing preference of the STAT complexes activated by IL-2 in CILL-2 cells is the same as for the complexes activated by GM-CSF in U937 cells. This is in accord with the fact that IL-2 and GM-CSF activate the same STAT protein, STAT5. The spacing preference for the STAT complexes activated by IL-13 in ME-180 cells was the same as for the STAT complexes activated by IL-4 in U-937 and ME-180 cells and is due to the likelihood that IL-4 and IL-13 both activate the same STAT protein, STAT6.

In NFS-60 cells, G-CSF activated two STAT complexes that were distinguishable by their differing mobilities in an EMSA. The slower-migrating complex comigrated with the STAT3 homodimer stimulated by IL-6 and was shown to

WO 95/28482

PCT/US95/04477

50

- contain STAT3 by antibody supershift experiments using a specific STAT3 antiserum. The faster-migrating complex contained an unidentified STAT complex that migrated like the STAT complexes activated by GM-CSF. In binding experiments, for the oligonucleotide series that were tested, the G-CSF-activated STAT3-like complex had spacing preferences indistinguishable from the STAT3 homodimer induced by IL-6. Similarly, the G-CSF-activated STAT complex that comigrated with the GM-CSF-activated STAT complexes had the same corresponding spacing preferences.

### EXAMPLE 3

- 10 Preparation of the reagents, cells, cell culture, nuclear extracts and Electrophoretic Mobility Shift Assays were performed as described in Example 1, with the exception that, in addition to some mentioned already in Example 1, the double-stranded probe oligonucleotides used in the Electrophoretic Mobility Shift Assays (EMSAs) were formed by annealing oligonucleotides with the sequences:

- 15 5'-GATCTGCTTCCCAAGAACGT -3' (SEQ ID NO. 136)  
3'-ACGAAGGGTCTTGCACTAG -5' (SEQ ID NO. 137)
- 5'-GATCCACTTCCCGGAACAGA -3' (SEQ ID NO. 138)  
20 3'-GTGAAGGGGCTTGCTCTAG -5' (SEQ ID NO. 139)
- 5'-GATCCACTTCCCAAGAACAGA -3' (SEQ ID NO. 140)  
3'-GTGAAGGGGCTTGCTCTAG -5' (SEQ ID NO. 141)
- 25 5'-GATCCACTTCCAGGAACAGA -3' (SEQ ID NO. 142)  
3'-GTGAAGGGTCTTGCTCTAG -5' (SEQ ID NO. 143)
- 5'-GATCTACTTCCCAAGAACATA -3' (SEQ ID NO. 144)  
3'-ATGAAGGGTCTTGTAATCTAG -5' (SEQ ID NO. 145)
- 30 5'-GATCCGCTTCCCAAGAACGGA -3' (SEQ ID NO. 146)  
3'-GCGAAGGGTCTTGCTCTAG -5' (SEQ ID NO. 147)
- 5'-GATCCACTTCTTAAGAACAGA -3' (SEQ ID NO. 148)  
35 3'-GTGAAGAATTCTTGCTCTAG -5' (SEQ ID NO. 149)
- 5'-GATCCACTTCCCAAGAACAGA -3' (SEQ ID NO. 150)  
3'-GTGAAGGGTCTTGCTCTAG -5' (SEQ ID NO. 151)

WO 95/28482

PCT/US95/04477

51

where the nucleotide sequences shown in bold type face correspond to nucleotide sequences, including their double-stranded complement, tested for activity as regulatory elements according to the present invention.

- Relative binding affinities, as determined from the EMSA results for oligonucleotides listed below, were visually rated and assigned according to the same visual rating scale as described in Example 1.

This visual rating system is sufficient to analyze distinguishable differences and trends in the EMSA binding data as opposed to specific numerical values. If desired, the use of a phosphor imager or densitometer (commercially available from e.g., Bio-Rad Laboratories) could provide a means to assess the differences described here quantitatively. Specific visual ratings of binding affinities for the regulatory elements of oligonucleotide SEQ ID NOs 58-59, 68-69, 70-71, 86-87, 100-101, 106-107, 108-109, 110-111, and 136-150 are shown in Table 14 below (in Table 14, bold type face highlight the core sequences).

Table 14: Relative EMSA binding affinities for a series of regulatory elements (core plus some flanking sequences) differing in flanking and spacing sequences to transcriptional regulatory proteins activated in response to the cytokines IL-4 and IL-13.

| SEQ ID | Regulatory Element | IL-4 | IL-13 |
|--------|--------------------|------|-------|
| 110    | CACTTCCCAAGAACAGA  | +++  | +++   |
| 70     | TGCTTCCCAGAACGT    | ++   | ++    |
| 86     | TGCTTCCCAGAACGT    | +    | +     |
| 136    | TGCTTCCCAAGAACGT   | ++   | ++    |
| 138    | CACTTCCCAGAACAGA   | ++   | ++    |
| 140    | CACTTCCCAGAACAGA   | ++   | ++    |
| 142    | CACTTCCCAGAACAGA   | ++   | ++    |

WO 95/28482

PCT/US95/04477

52

|     |                   |    |      |
|-----|-------------------|----|------|
| 144 | TACTTCCCAAGAACATA | ++ | ++   |
| 146 | CGCTTCCCAAGAACGGA | ++ | ++   |
| 148 | CACTTCTTAAGAACAGA | ++ | ++   |
| 150 | CACTTCCCAAGAACAGA | ++ | ++   |
| 58  | TGCTTCCCGAACGT    | ++ | n.d. |
| 68  | TGCTTCCCGAACGT    | ++ | ++   |
| 100 | TGCTTCTCAGAACGT   | ++ | ++   |
| 106 | GATTTCCTCGAAATG   | ++ | ++   |
| 108 | CAATTCTAAGAAAGGA  | ++ | n.d. |

n.d. = not determined

The data in Table 13 show that the IL-4 and IL-13-activated STAT complexes can bind to all of the listed sequences with similar affinity (with the exception of SEQ ID NO. 86, which was slightly lower in affinity). These data are fully consistent with the binding data described in Example 1 herein.

#### Transient transfection assays

The reporter plasmids SEQID58x4-TK-LUC, SEQID68x4-TK-LUC, SEQID70x4-TK-LUC, SEQID100x4-TK-LUC, SEQID106x4-TK-LUC, SEQID108x4-TK-LUC, SEQID110x4-TK-LUC, SEQID86x4-TK-LUC, SEQID136x4-TK-LUC, SEQID138x4-TK-LUC, SEQID140x4-TK-LUC, SEQID142x4-TK-LUC, SEQID144x4-TK-LUC, SEQID146x4-TK-LUC, SEQID148x4-TK-LUC, and SEQID150x4-TK-LUC contain four copies of oligonucleotide sequences with the indicated SEQ ID NOs linked to the promoter of the Herpes Simplex virus thymidine kinase gene at position -35 with respect to the cap site. See FIG. 1. The reference reporter, TK-LUC (P. Lamb et al., 83 *Blood* 2063 (1994)), the disclosure of which is

WO 95/28482

PCT/US95/04477

53

herein incorporated by reference, contains the simple Herpes Simplex virus thymidine kinase promoter with no regulatory element. These chimeric promoters drive the expression of the structural gene for firefly luciferase.

- ME-180 cells were transfected with the reporter plasmids of above by calcium phosphate coprecipitation as described in Example 1. The results of these transfections are shown below in Tables 15 and 16. Numbers given are the mean fold inductions ('fold induction' is defined as the normalized response in a cytokine-treated sample divided by the normalized response in an untreated sample).
- Table 15: Transcriptional induction in ME-180 cells of reporter constructs incorporating IL-4/IL-13-selective regulatory elements showing effect of flanking and spacing sequence (core sequence highlighted). The values given are mean fold inductions in response to the indicated cytokine. The value in the parentheses is the number of experiments included to calculate the mean.

| Reporter         | Core Element       | IL-4    | IL-13   |
|------------------|--------------------|---------|---------|
| TK-LUC           | none               | 0.9 (3) | 1.0 (3) |
| SEQID110x4TK-LUC | CACTTCCCAAGAACAGA  | 22 (3)  | 9.7 (3) |
| SEQID70x4TK-LUC  | TGCTTCCCCGGAACGT   | 1.3 (3) | 1.0 (3) |
| SEQID86x4TK-LUC  | TGCTTCCCCAGAACGT   | 1.2 (3) | 1.1 (3) |
| SEQID136x4TK-LUC | TGCTTCCCAAGAACGT   | 1.5 (3) | 1.2 (3) |
| SEQID138x4TK-LUC | CACTTCCCCGGAACAGA  | 2.7 (3) | 1.6 (3) |
| SEQID140x4TK-LUC | CACTTCCCCAGAACAGA  | 8.0 (3) | 3.0 (3) |
| SEQID142x4TK-LUC | CACTTCCCCAGGAACAGA | 10 (3)  | 6.4 (3) |
| SEQID144x4TK-LUC | TACTTCCCAAGAACATA  | 3.0 (3) | 1.5 (3) |
| SEQID146x4TK-LUC | CGCTTCCCAAGAACGGA  | 1.5 (3) | 1.3 (3) |

WO 95/28482

PCT/US95/04477

54

|                  |                   |         |         |
|------------------|-------------------|---------|---------|
| SEQID148x4TK-LUC | CACTTCTTAAGAACAGA | 7.3 (3) | 3.3 (3) |
| SEQID150x4TK-LUC | CACTTCCAAGAACAGA  | 1.2 (3) | 1.1 (3) |

Table 16: Regulatory elements with a five base pair spacing that bind to the IL-4-Activated STAT complexes do not mediate transcriptional induction in ME-180 cells treated with IL-4 (spacing nucleotides highlighted). The values given are mean fold inductions in response to the indicated cytokine. The value in the parentheses is the number of experiments included to calculate the mean.

| Reporter         | Core Element     | IL-4    |
|------------------|------------------|---------|
| SEQID58x4TK-LUC  | TGCTTCCCCGAACGT  | 1.2 (3) |
| SEQID68x4TK-LUC  | TGCTTCCCGAACGT   | 1.1 (3) |
| SEQID100x4TK-LUC | TGCTTCTCAGAACGT  | 1.2 (3) |
| SEQID106x4TK-LUC | GATTTCCCGAAATG   | 0.8 (3) |
| SEQID108x4TK-LUC | CAATTCTAAGAAAGGA | 0.8 (3) |
| SEQID150x4TK-LUC | CACTTCCAAGAACAGA | 1.2 (3) |

The data summarized in Table 15 demonstrate that the identity of the spacing nucleotides and flanking nucleotides influence whether or not a given regulatory sequence will be transcriptionally active (defined herein as greater than a two fold induction). In general, the responses in ME-180 cells to IL-13 were lower than to IL-4, probably reflecting a lower number of IL-13 receptors on ME-180 cells (this correlates with less STAT complex being activated by IL-13 as compared to IL-4). The effect of the flanking nucleotides is shown especially well by comparing the SEQ ID NO. 140x4TK-LUC entry with the SEQ ID NO. 70x4TK-LUC entry. SEQ ID NO. 140x4TK-LUC is activated by IL-4, while SEQ ID NO. 70x4TK-LUC is not, and the only difference between SEQ ID NO. 140 and SEQ ID NO. 70 are the two nucleotides



on the 5' and 3' flanks of the core element. The data summarized in Table 16 demonstrate that for all of the tested response elements with a three base pair spacing, though they bind well to the IL-4-activated STAT complexes, they surprisingly do not mediate a transcriptional response in ME-180 cells in the context of the TK minimal promoter. This is true even if the flanking nucleotides are "optimal" (taken from SEQ ID NO. 110) as shown in Tables 15 and 16 for the entries for SEQ ID NO. 150x4TK-LUC. Thus, although all of the sequences shown in Tables 15 and 16 are capable of binding the STAT complexes activated by IL-4 and IL-13 (shown in Table 14), not all mediate a transcriptional induction in response to IL-4 or IL-13 (defined as greater than two fold induction). Nevertheless, because the ability of a regulatory element to bind to a given STAT complex is a prerequisite for that sequence to be transcriptionally active, the observed spacing preferences are thus an important determinant of selective transcriptional activation.

From the data in Table 15, it is possible to identify a general sequence that will not only bind the IL-4 and IL-13 activated STAT complexes selectively, but, importantly, will also mediate a transcriptional induction in response to IL-4 and IL-13. That general sequence is: 5' - ANTTCCNNNGAANA - 3' (SEQ ID NO. 152) [and its double-stranded complement: 5' - TTTCCNNNGAANT - 3' (SEQ ID NO. 153)], where N is a nucleotide independently selected from A, T, C or G. This sequence includes nucleotides outside of the core element that, from the data in Table 15, have been shown to influence transcriptional activation. Accordingly, preferred oligonucleotide sequences selective for IL-4 and IL-13 are selected from the group consisting of ACTTCCAAGAACA (SEQ ID NO. 154), TGTCTTGGGAAGT (SEQ ID NO. 155), ACTTCCCGGAACA (SEQ ID NO. 156), TGTTCGGGAAGT (SEQ ID NO. 157), ACTTCCCCAGAACA (SEQ ID NO. 158), TGTTCGGGAAGT (SEQ ID NO. 159), ACTTCCAGGAACA (SEQ ID NO. 160), TGTCTTGGGAAGT (SEQ ID NO. 161), ACTTCCAAGAACA (SEQ ID NO. 162), TGTCTTGGGAAGT (SEQ ID NO. 163), ACTTCTAAGAACA (SEQ ID NO. 164), and TGTTCCTAAGAAGT (SEQ ID NO. 165).

WO 95/28482

PCT/US95/04477

56

While in accordance with the patent statutes, description of the preferred weight fractions, and processing conditions have been provided, the scope of the invention is not to be limited thereto or thereby. Various modifications and alterations of the present invention will be apparent to those skilled in the art without departing  
5 from the scope and spirit of the present invention.

Consequently, for an understanding of the scope of the present invention, reference is made to the following claims.

WO 95/28482

57

PCT/US95/04477

## SEQUENCE LISTING

5

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- 10 (A) NAME: LIGAND PHARMACEUTICALS INCORPORATED  
(B) STREET: 9393 TOWNE CENTRE DRIVE  
(C) CITY: SAN DIEGO  
(D) STATE: CALIFORNIA  
(E) COUNTRY: US  
15 (F) POSTAL CODE (ZIP): 92121  
(G) TELEPHONE: (619) 550-7675  
(H) TELEFAX: (619) 535-3906

- 20 (ii) TITLE OF INVENTION: DNA SPACER REGULATORY ELEMENTS  
RESPONSIVE TO CYTOKINES AND METHODS FOR THEIR USE

(iii) NUMBER OF SEQUENCES: 165

## (iv) COMPUTER READABLE FORM:

- 25 (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
30 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (vi) PRIOR APPLICATION DATA:

- 35 (A) APPLICATION NUMBER: US 08/228,935  
(B) FILING DATE: 14-APR-1994

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 8 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,

WO 95/28482

58

PCT/US95/04477

## SYNTHETIC DNA\*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

5 TTNNNNNAA

8

(2) INFORMATION FOR SEQ ID NO: 2:

10 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TTNNNNNAA

9

25 (2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

35

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TTNNNNNAA

10

45

(2) INFORMATION FOR SEQ ID NO: 4:

WO 95/28482

59

PCT/US95/04477

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TTNNNNNNA A

11

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TTCNNGAA

8

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

WO 95/28482

60

PCT/US95/04477

TTCNNNGAA

9

## 5 (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- 15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

20 TTCNNNNGAA

10

## (2) INFORMATION FOR SEQ ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 8 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- 30 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TTANNGAA

8

## 40 (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

61

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTANNNGAA

9

10

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

25 TTANNNGAA

10

(2) INFORMATION FOR SEQ ID NO: 11:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TTANNNNNGA A

11

45 (2) INFORMATION FOR SEQ ID NO: 12:

WO 95/28482

62

PCT/US95/04477

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
TTANNNTTA 9
- (2) INFORMATION FOR SEQ ID NO: 13:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"



WO 95/28482

63

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TTANNNNNTAA

10

5

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

20 TTANNNNNTA A

11

(2) INFORMATION FOR SEQ ID NO: 15:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TTTNNGAA

8

40 (2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 base pairs

(B) TYPE: nucleic acid

45

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

WO 95/28482

64

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TTTNNNGAA

9

10

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

25 TTTNNNGAA

10

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

WO 95/28482

65

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TTTNNNNNGA A

11

5

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- 15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

20 TTTNNNTAA

9

(2) INFORMATION FOR SEQ ID NO: 20:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TTTNNNTAA

10

40 (2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 11 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

66

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

TTTNNNNNTA A

11

10

- (2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 8 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

20

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

25 TTCCCGAA

8

- (2) INFORMATION FOR SEQ ID NO: 23:

30

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

- (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

67

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

TTCCCCGAA

9

5

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

20 TTCCCCGAA

10

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TTCCGGAA

8

40

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45

WO 95/28482

68

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

TTCCCGGAA

9

10

- (2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

20

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

25

TTCCCGGAA

10

- (2) INFORMATION FOR SEQ ID NO: 28:

30

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

- (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

69

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

TTCCTGGAA

9

5

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

20 TTCCTGGAA

10

(2) INFORMATION FOR SEQ ID NO: 30:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

TTCCAGAA

8

40 (2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 base pairs

(B) TYPE: nucleic acid

45

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

WO 95/28482

70

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

TTCCCAGAA

9

10

- (2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

20

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

25 TTCCCCAGAA

10

- (2) INFORMATION FOR SEQ ID NO: 33:

30

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

- (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"



WO 95/28482

71

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

TTCTTTGAA

9

5

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

20 TTCTTTTGAA

10

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

TTTCAGAA

9

40 (2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

WO 95/28482

72

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

TTCTCCAGAA

10

10

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

25

TTACCGTAA

9

(2) INFORMATION FOR SEQ ID NO: 38:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

73

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

TTACCCGTAA

10

5

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

20 TTACCCGTAA

11

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

TTCCCGTAA

9

40 (2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45

WO 95/28482

74

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

TTCCCCGTAA

10

10

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 11 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

25

TTCCCCGTAA

11

(2) INFORMATION FOR SEQ ID NO: 43:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

75

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

TTCTGTAA

8

5

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- 15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

20 TTCTCGTAA

9

(2) INFORMATION FOR SEQ ID NO: 45:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

TTCTCCGTAA

10

40 (2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 11 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

76

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

TTCTCCCGTA A

11

10

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

TTCCCAAGAA

10

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

WO 95/28482

77

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

TTTCCCGTAA

10

5

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

20 TTCCCAGGAA

10

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

TTCTTAAGAA

10

40 (2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45

WO 95/28482

78

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

TTCTAAGAA

9

10

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

25 TTTCCCCGAA

10

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35



WO 95/28482

79

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

TTTCTAAGAA

10

5

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

20 GATCTGCTTC CGAACGT

17

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: other nucleic acid

30

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

35

GATCACGTTC GGAAGCA

17

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

WO 95/28482

80

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

GATCTGCTTC CCGAACGT

18

10

- (2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

25 GATCACGTTC GGAAGCA

18

- (2) INFORMATION FOR SEQ ID NO: 58:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

35

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

81

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

GATCTGCTTC CCCGAACGT

19

5

(2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- 15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

20 GATCACGTTT GGGGAAGCA

19

(2) INFORMATION FOR SEQ ID NO: 60:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

GATCTGCTTC CCCCGAACGT

20

40 (2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

82

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

GATCACGTTT GGGGAAGCA

20

10

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

15

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

25 GATCTGCTTC CCCC GAACG T

21

(2) INFORMATION FOR SEQ ID NO: 63:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

83

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

GATCACGTTT GGGGGGAAGC A

21

5

(2) INFORMATION FOR SEQ ID NO: 64:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

20 GATCTGCTTC CCCCCGAAC GT

22

(2) INFORMATION FOR SEQ ID NO: 65:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

GATCACGTTT GGGGGGAAG CA

22

40 (2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45

WO 95/28482

84

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

GATCTGCTTC CGGAACGT

18

10

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

15

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

25

GATCACGTTT CGGAAGCA

18

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

30

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

85

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

GATCTGCTTC CCGGAACGT

19

5

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- 15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

20 GATCACGTTT CCGGAAGCA

19

(2) INFORMATION FOR SEQ ID NO: 70:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

GATCTGCTTC CCCGGAACGT

20

40 (2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

86

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

GATCACGTTC CGGGGAAGCA

20

10

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

15

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

25 GATCTGCTTC CCCCGBAACG T

21

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

30

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"



WO 95/28482

87

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

GATCACG TTC CGGGGGAAGC A

21

5

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

20 GATCTGCTTC CCCCCGGAAC GT

22

(2) INFORMATION FOR SEQ ID NO: 75:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

GATCACG TTC CGGGGGAAG CA

22

40 (2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45

WO 95/28482

88

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

GATCTGCTTC CTGGAACGT

19

10

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

25 GATCACGTTC CAGGAAGCA

19

(2) INFORMATION FOR SEQ ID NO: 78:

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

89

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

GATCTGCTTC CTTGAACGT

20

5

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- 15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

20 GATCAGGTTT CAAGGAAGCA

20

(2) INFORMATION FOR SEQ ID NO: 80:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

GATCTGCTTC CTTTGAACG T

21

40 (2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
45 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

90

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

GATCACGTTTCAAAGGAAGC A

21

10

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

25 GATCTGCTTCAGAACGT

18

(2) INFORMATION FOR SEQ ID NO: 83:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

91

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

GATCACGTTT TGGAAGCA

18

5

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

15

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

20 GATCTGCTTC CCAGAACGT

19

(2) INFORMATION FOR SEQ ID NO: 85:

25

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

GATCACGTTT TGGGAAGCA

19

40 (2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

WO 95/28482

92

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

GATCTGCTTC CCCAGAACGT

20

10

- (2) INFORMATION FOR SEQ ID NO: 87:

- (i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

20

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

25 GATCAGTTC TGGGGAAGCA

20

- (2) INFORMATION FOR SEQ ID NO: 88:

30

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

- (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

93

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

GATCTGCTTC CCCAGAACG T

21

5

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- 15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

20 GATCAGGTC TGGGGAAGC A

21

(2) INFORMATION FOR SEQ ID NO: 90:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

GATCTGCTTC TTGAACGT

18

40 (2) INFORMATION FOR SEQ ID NO: 91:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

94

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

GATCACGTTT AAGAAGCA

18

10

(2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

25 GATCTGCTTC TTGAAAGT

19

(2) INFORMATION FOR SEQ ID NO: 93:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"



WO 95/28482

95

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

GATCACGTTT AAAGAAGCA

19

5

(2) INFORMATION FOR SEQ ID NO: 94:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

20 GATCTGCTTC TTTTGAACGT

20

(2) INFORMATION FOR SEQ ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xc) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

35

GATCACGTTT AAAAGAAGCA

20

40 (2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45

WO 95/28482

96

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

GATCTGCTTC TTTTGAACG T

21

10

(2) INFORMATION FOR SEQ ID NO: 97:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

25 GATCACGTTT AAAAAGAAGC A

21

(2) INFORMATION FOR SEQ ID NO: 98:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

35

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

97

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

GATCTGCTTC TAGAACGT

18

5

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

20 GATCAGGTC TAGAAGCA

18

(2) INFORMATION FOR SEQ ID NO: 100:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

GATCTGCTTC TCAGAACGT

19

40

(2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45

WO 95/28482

98

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

GATCACGTTT TGAGAAGCA

19

10

(2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

15

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

25

GATCTGCTTC TCCAGAACGT

20

(2) INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

30

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

99

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

GATCACGTTT TGGAGAAGCA

20

5

(2) INFORMATION FOR SEQ ID NO: 104:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

15

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID, SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

20 GATCTGCTTC TCCAGAACG T

21

(2) INFORMATION FOR SEQ ID NO: 105:

25

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID, SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

GATCACGTTT TGGGAGAAGC A

21

40

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

WO 95/28482

100

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

GATCGATTTC CCCGAAATG

19

10

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

15

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

25 GATCCATTTC GGGGAAATC

19

(2) INFORMATION FOR SEQ ID NO: 108:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

101

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

GATCCAATTT CTAAGAAAGG A

21

5

(2) INFORMATION FOR SEQ ID NO: 109:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

15

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID, SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

20 GATCTCCTTT CTTAGAAATT G

21

(2) INFORMATION FOR SEQ ID NO: 110:

25

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID, SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

GATCCAATTC CCAAGAACAG A

21

40

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

WO 95/28482

102

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

GATCTCTGTT CTTGGGAAGT G

21

10

(2) INFORMATION FOR SEQ ID NO: 112:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

25 GATCTGCTTA CGTAACGT

18

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35



WO 95/28482

103

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

GATCACGTTA CGTAAGCA

18

5

(2) INFORMATION FOR SEQ ID NO: 114:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

15

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

20 GATCTGCTTA CCGTAACGT

19

(2) INFORMATION FOR SEQ ID NO: 115:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

GATCACGTTA CCGTAAGCA

19

40 (2) INFORMATION FOR SEQ ID NO: 116:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45

WO 95/28482

104

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

GATCTGCTTA CCCGTAACGT

20

10

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid

20

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

25

GATCACGTTA CGGGTAAGCA

20

(2) INFORMATION FOR SEQ ID NO: 118:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

35

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

WO 95/28482

105

PCT/US95/04477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

GATCTGCTTA CCCGTAACG T

21

5 (2) INFORMATION FOR SEQ ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

15 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

GATCACGTTA CGGGGTAAGC A

21

(2) INFORMATION FOR SEQ ID NO: 120:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

30 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

35 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

40 GATCTGCTTC CGTAACGT

18

(2) INFORMATION FOR SEQ ID NO: 121:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

WO 95/28482

106

PCT/US95/04477

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 5 (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

10 GATCACGTTA CGGAAGCA 18

(2) INFORMATION FOR SEQ ID NO: 122:

- 15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
20 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

- 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

GATCTGCTTC CCGTAACGT 19

- 30 (2) INFORMATION FOR SEQ ID NO: 123:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: linear

WO 95/28482

107

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

GATCACGTTA CGGGAAGCA

19

10 (2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

25 GATCTGCTTC CCCGTAACGT

20

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

40 GATCACGTTA CGGGGAAGCA

20

WO 95/28482

108

PCT/US95/04477

## (2) INFORMATION FOR SEQ ID NO: 126:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

GATCTGCTTC CCCCGTAACG T

21

## (2) INFORMATION FOR SEQ ID NO: 127:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

GATCACGTTA CGGGGAAGC A

21

## (2) INFORMATION FOR SEQ ID NO: 128:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

109

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

GATCTGCTTC TGTAACGT

18

10 (2) INFORMATION FOR SEQ ID NO: 129:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

25 GATCACGTTA CAGAAGCA

18

(2) INFORMATION FOR SEQ ID NO: 130:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

GATCTGCTTC TCGTAACGT

19

45 (2) INFORMATION FOR SEQ ID NO: 131:

(i) SEQUENCE CHARACTERISTICS:

WO 95/28482

110

PCT/US95/04477

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID, SYNTHETIC DNA"

- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

GATCACGTTA CGAGAAGCA

19

- 15 (2) INFORMATION FOR SEQ ID NO: 132:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID, SYNTHETIC DNA"

25

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

GATCTGCTTC TCCGTAACGT

20

30

- (2) INFORMATION FOR SEQ ID NO: 133:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35



WO 95/28482

111

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

GATCACGTGA CGGAGAAGCA

20

10 (2) INFORMATION FOR SEQ ID NO: 134:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

20 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

GATCTGCTTC TCCCGTAACG T

21

25

(2) INFORMATION FOR SEQ ID NO: 135:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

GATCACGTGA CGGAGAAGC A

21

40

45 (2) INFORMATION FOR SEQ ID NO: 136:

WO 95/28482

112

PCT/US95/04477

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

GATCTGCTTC CCAAGAACGT

## (2) INFORMATION FOR SEQ ID NO: 137:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

GATCACGTTC TTGGGAAGCA

## (2) INFORMATION FOR SEQ ID NO: 138:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

WO 95/28482

113

PCT/US95/04477

GATCCACTTC CCCGGAACAG A

21

(2) INFORMATION FOR SEQ ID NO: 139:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

10

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

GATCTCTGTT CCGGGAAGT G

21

20

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

25

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

30

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

35

GATCCACTTC CCCAGAACAG A

21

WO 95/28482

114

PCT/US95/04477

## (2) INFORMATION FOR SEQ ID NO: 141:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- 10 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

15 GATCTCTGTT CTGGGGAAGT G 21

## (2) INFORMATION FOR SEQ ID NO: 142:

## (i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- 25 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

30 GATCCACTTC CCAGGAACAG A 21

## 35 (2) INFORMATION FOR SEQ ID NO: 143:

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

115

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

GATCTCTGTT CCTGGGAAGT G

21

10 (2) INFORMATION FOR SEQ ID NO: 144:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

25 GATCTACTTC CCAAGAACAT A

21

(2) INFORMATION FOR SEQ ID NO: 145:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:

40

GATCTATGTT CTTGGGAAGT A

21

(2) INFORMATION FOR SEQ ID NO: 146:

45

(i) SEQUENCE CHARACTERISTICS:

WO 95/28482

116

PCT/US95/04477

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

GATCCGCTTC CCAAGAACGG A

21

15 (2) INFORMATION FOR SEQ ID NO: 147:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

GATCTCCGTT CTGGGAAGC G

30

21

(2) INFORMATION FOR SEQ ID NO: 148:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

WO 95/28482

117

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

GATCCACTTC TTAAGAACAG A

21

10 (2) INFORMATION FOR SEQ ID NO: 149:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
15 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

GATCTCTGTT CTTAAGAAGT G

21

25

(2) INFORMATION FOR SEQ ID NO: 150:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
30 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

GATCCACTTT CCAAGAACAG A

21

40

(2) INFORMATION FOR SEQ ID NO: 151:

45

- (i) SEQUENCE CHARACTERISTICS:

WO 95/28482

118

PCT/US95/04477

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5

- (ii) MOLECULE TYPE: other nucleic acid
- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

GATCTCTGTT CTGGAAAGT G

21

- (2) INFORMATION FOR SEQ ID NO: 152:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

- (ii) MOLECULE TYPE: other nucleic acid
- (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

25

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

ATTTCNNNG AANA

14

30

- (2) INFORMATION FOR SEQ ID NO: 153:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35



WO 95/28482

119

PCT/US95/04477

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153:

TNTTCNNNG AANT

14

10 (2) INFORMATION FOR SEQ ID NO: 154:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

25 ACTTCCAAG AACA

14

(2) INFORMATION FOR SEQ ID NO: 155:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

40 TGTTCTGGG AAGT

14

(2) INFORMATION FOR SEQ ID NO: 156:

45

- (i) SEQUENCE CHARACTERISTICS:

WO 95/28482

120

PCT/US95/04477

- (A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- 5 (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:  
ACTTCCCCGG AACA 14
- 15 (2) INFORMATION FOR SEQ ID NO: 157:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
20 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
25 SYNTHETIC DNA"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:  
TGTTCCGGGG AAGT 14
- 30 (2) INFORMATION FOR SEQ ID NO: 158:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
35 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

121

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

ACTTCCCAG AACA

14

10 (2) INFORMATION FOR SEQ ID NO: 159:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20 (A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

25 TGTTCGGGG AAGT

14

(2) INFORMATION FOR SEQ ID NO: 160:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

40

ACTTCCCAGG AACA

14

(2) INFORMATION FOR SEQ ID NO: 161:

45

(i) SEQUENCE CHARACTERISTICS:

WO 95/28482

122

PCT/US95/04477

- (A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- 5 (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:  
TGTTCTCTGGG AAGT 14

## 15 (2) INFORMATION FOR SEQ ID NO: 162:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
20 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
25 SYNTHETIC DNA"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:  
30 ACTTCCCAAG AACA 14

## (2) INFORMATION FOR SEQ ID NO: 163:

- (i) SEQUENCE CHARACTERISTICS:  
35 (A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

WO 95/28482

123

PCT/US95/04477

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
SYNTHETIC DNA"

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

TGTICTTGGG AAGT

14

10 (2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
15 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
20 SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

25 ACTTCTTAAG AACA

14

(2) INFORMATION FOR SEQ ID NO: 165:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 base pairs  
30 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "OTHER NUCLEIC ACID,  
35 SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

40 TGTICTTAAG AAGT

14

WO 95/28482

124

PCT/US95/04477

What is claimed is:

1. An oligonucleotide sequence comprising a regulatory element of the nucleotide sequence TTN<sub>x</sub>AA, wherein N is independently selected from A, T, C or G and x is 4, 5, 6 or 7, and wherein the regulatory element binds to an activated transcriptional regulatory protein in response to a signaling molecule.
2. An oligonucleotide sequence according to claim 1, wherein the signaling molecule comprises a polypeptide, an oligosaccharide, a non-peptidyl molecule or synthetic mimics thereof.
3. An oligonucleotide sequence according to claim 2, wherein the polypeptide comprises a cytokine, a hormone, an antibody, a cell-surface antigen or synthetic mimics thereof.
4. An oligonucleotide sequence according to claim 1, wherein the transcriptional regulatory protein comprises a STAT protein.
5. An oligonucleotide sequence according to claim 4, wherein the STAT protein is selected from the group consisting of STAT1 $\alpha$  protein, STAT1 $\beta$  protein, STAT2 protein, STAT3 protein, STAT4 protein, STAT5 protein and STAT6 protein.
6. An oligonucleotide sequence according to claim 1, wherein the nucleotide sequence is selected from the group consisting of TTCN<sub>y</sub>GAA, TTAN<sub>y</sub>GAA, TTAN<sub>y</sub>TAA, TTTN<sub>y</sub>GAA and TTTN<sub>y</sub>TAA, where N is independently selected from A, T, C or G and y is 2, 3, 4 or 5.
7. An oligonucleotide sequence according to claim 1 selected from the group consisting of TTCCCGAA (SEQ ID NO. 22), TTCCCGAA (SEQ ID NO.

1-  $\rightarrow$  "half site" selected from  
5' and 3'  
2- which strand  
half site  
is separated  
by  
does it still  
bind monomers  
is nucleotides  
3- where  
the site of  
claim 2  
consists of  
2 half sites  
"organized" in  
an inserted  
repeat  
configuration

WO 95/28482

PCT/US95/04477

125

23), TTCCCCCGAA (SEQ ID NO. 24), TTCCCGAA (SEQ ID NO. 25),  
 TTCCCCGAA (SEQ ID NO. 26), TTCCCCCGAA (SEQ ID NO. 27), TTCCTGGAA  
 (SEQ ID NO. 28), TTCCTTGGAA (SEQ ID NO. 29), TTCCAGAA (SEQ ID NO. 30),  
 TTCCCAGAA (SEQ ID NO. 31), TTCCCCAGAA (SEQ ID NO. 32), TTCTTTGAA  
 5 (SEQ ID NO. 33), TTCTTTTGGAA (SEQ ID NO. 34), TTCTCAGAA (SEQ ID NO.  
 35), TTCTCCAGAA (SEQ ID NO. 36), TTACCGTAA (SEQ ID NO. 37),  
 TTACCCGTAA (SEQ ID NO. 38), TTACCCCGTAA (SEQ ID NO. 39),  
 TTCCCGTAA (SEQ ID NO. 40), TTCCCGGTAA (SEQ ID NO. 41),  
 TTCCCCCGTAA (SEQ ID NO. 42), TTCTGTAA (SEQ ID NO. 43), TTCTCGTAA  
 10 (SEQ ID NO. 44), TTCTCCGTAA (SEQ ID NO. 45), TTCTCCCGTAA (SEQ ID NO.  
 46), TTCCAAGAA (SEQ ID NO. 47), TTCCCGTAA (SEQ ID NO. 48),  
 TTCCCAGGAA (SEQ ID NO. 49), TTCTTAAGAA (SEQ ID NO. 50), TTCTAAGAA  
 (SEQ ID NO. 51), TTTCCCCGAA (SEQ ID NO. 52), and TTTCTAAGAA (SEQ ID  
 NO. 53).

15

8. An oligonucleotide sequence according to claim 6 comprising  
 $TTCN_yGAA$  or  $TTAN_yGAA$ , wherein  $y$  is 2 and the regulatory element selectively  
 binds to a cytokine-activated transcriptional regulatory protein comprising STAT3  
 protein.

20

9. An oligonucleotide sequence according to claim 8, wherein the  
 cytokine is selected from the group consisting of IL-6, LIF, EGF, PDGF, G-CSF, IL-10,  
 IL-11, growth hormone and Oncostatin M.

25

10. An oligonucleotide sequence according to claim 6 comprising  
 $TTAN_yGAA$  or  $TTAN_yTAA$ , wherein  $y$  is 3 and the regulatory element binds to a  
 cytokine-activated transcriptional regulatory protein comprising STAT1 $\alpha$  protein  
 and/or STAT3 protein.

WO 95/28482

PCT/US95/04477

126

11. An oligonucleotide sequence according to claim 10, wherein the cytokine is selected from the group consisting of IL-6, LIF, EGF, PDGF, G-CSF, IL-10, IL-11, IFN $\gamma$ , growth hormone and Oncostatin M.
- 5 12. An oligonucleotide sequence according to claim 6 comprising TTCN $\gamma$ GAA, wherein  $\gamma$  is 3 and the regulatory element binds to a cytokine-activated transcriptional regulatory protein comprising STAT1 protein, STAT3 protein, STAT4 protein, STAT5 protein and/or STAT6 protein.
- 10 13. An oligonucleotide sequence according to claim 12, wherein the cytokine is selected from the group consisting of IL-6, LIF, EGF, PDGF, G-CSF, IL-10, IL-11, growth hormone, Oncostatin M, GM-CSF, IL-2, IL-3, IL-4, IL-5, IL-7, IL-9, IL-12, IL-13, IL-15, Tpo, Epo, prolactin and IFN $\gamma$ .
- 15 14. An oligonucleotide sequence according to claim 6 comprising TTCTTTGAA, TTCTCAGAA, TTAN $\gamma$ GAA or TTAN $\gamma$ TAA, wherein  $\gamma$  is 4 and the regulatory element selectively binds to a cytokine-activated transcriptional regulatory protein comprising STAT5 protein and/or STAT6 protein.
- 20 15. An oligonucleotide sequence according to claim 14, wherein the cytokine is selected from the group consisting of IL-2, IL-3, IL-4, IL-5, IL-7, IL-9, IL-13, IL-15 GM-CSF, G-CSF, Tpo, Epo, growth hormone and prolactin.
- 25 16. An oligonucleotide sequence according to claim 6 comprising TTCN $\gamma$ GAA, wherein  $\gamma$  is 4 and the regulatory element selectively binds to a cytokine-activated transcriptional regulatory protein comprising STAT6 protein.
17. An oligonucleotide sequence according to claim 16, wherein the cytokine is selected from the group consisting of IL-4, IL-7, IL-9, IL-13 and IL-15.



WO 95/28482

PCT/US95/04477

127

18. An oligonucleotide sequence according to claim 6 comprising  
TTAN<sub>y</sub>GAA or TTAN<sub>y</sub>TAA, wherein y is 5 and the regulatory element selectively  
binds to a cytokine-activated transcriptional regulatory protein comprising STAT6  
5 protein.
19. An oligonucleotide sequence according to claim 18, wherein the  
cytokine is selected from the group consisting of IL-4, IL-7, IL-9, IL-13 and IL-15.
20. An oligonucleotide sequence according to claim 1 which is double  
10 stranded.
21. A DNA construct comprising the oligonucleotide sequence of  
claim 1 operably linked to a promoter, which promoter is operably linked to a  
15 heterologous gene, wherein the DNA construct is linked in such a manner that the gene  
is under the transcriptional control of the oligonucleotide sequence and promoter.
22. A DNA construct according to claim 21, wherein the promoter is  
selected from the group consisting of the gene promoter of the Herpes simplex virus  
20 Thymidine kinase, the adenovirus Elb and yeast alcohol dehydrogenase.
23. A DNA construct according to claim 21, wherein the heterologous  
gene comprises the gene for luciferase, chloramphenicol acetyl transferase,  $\beta$ -  
galactosidase, secreted placental alkaline phosphatase, human growth hormone, t-PA,  
25 green fluorescent protein or interferon.
24. A DNA construct according to claim 21, wherein the  
oligonucleotide sequence comprises a multimer of at least one of the regulatory  
elements of claim 1.

30

WO 95/28482

PCT/US95/04477

128

25. A DNA construct according to claim 21, wherein the promoter comprises the gene promoter of the Herpes simplex virus thymidine kinase promoter and the heterologous gene comprises the gene for luciferase.
- 5 26. A DNA construct according to claim 21, wherein the oligonucleotide sequence is selected from the group consisting of TTCCCGAA (SEQ ID NO. 22), TTCCCGAA (SEQ ID NO. 26), TTCCCGAA (SEQ ID NO. 23), TTCCCGAA (SEQ ID NO. 27), TTCCAAGAA (SEQ ID NO. 47), TTCTAAGAA (SEQ ID NO. 51), TTCCCAGAA (SEQ ID NO. 32),  
10 TTCCCAGAA (SEQ ID NO. 49), TTCTAAGAA (SEQ ID NO. 50), TTCCCGAA (SEQ ID NO. 52), and TTTCTAAGAA (SEQ ID NO. 53).
27. A host cell transfected with the DNA construct of claim 21.
- 15 28. A host cell according to claim 27, wherein the cell comprises a HepG2 cell, U937 cell, ME-180 cell, a TF-1 cell or NFS-60 cell.
29. A method for the controlled expression of a heterologous gene of interest comprising culturing a host cell according to claim 27 comprising appropriate  
20 transcriptional regulatory proteins in the presence of a signaling molecule.
30. A method for controlled expression according to claim 29, wherein the signaling molecule comprises a cytokine and the transcriptional regulatory protein comprises a STAT protein.
- 25

128

SUBSTITUTE SHEET (RULE 26)

WO 95/28482

PCT/US95/04477

129

31. A method for detecting the presence of a transcriptional regulatory protein in a sample comprising contacting the sample with an oligonucleotide sequence according to claim 1 under conditions where the transcriptional regulatory protein is activated and binds with the oligonucleotide sequence to form a complex, and detecting  
5 the presence of the complex in the sample.

32. A method according to claim 31, further comprising, separating the complex from the sample and the oligonucleotide sequence from the complex to yield the transcriptional regulatory protein.

10

33. A method for measuring the ability of a compound to act as an agonist of gene transcription comprising:

(a) contacting the compound with a host cell according to claim 27 under conditions in which the heterologous gene is capable of being expressed in  
15 response to the compound; and

(b) comparing the level of gene expression in step (a) with the level of gene expression from the host cell in the absence of the compound.

34. A method for measuring the ability of a compound to act as an  
20 antagonist of gene transcription comprising:

(a) contacting the compound with a host cell according to claim 27 in the presence of a predetermined amount of a signaling molecule under conditions in which the heterologous gene is capable of being expressed in response to the signaling molecule; and

25 (b) comparing the level of gene expression in step (a) with the level of gene expression from the host cell in the presence of the signaling molecule, but the absence of the compound.

129

SUBSTITUTE SHEET (RULE 28)

WO 95/28482

PCT/US95/04477

130

35. A method according to claim 33, wherein the heterologous gene comprises the gene for luciferase, chloramphenicol acetyl transferase,  $\beta$ -galactosidase, green fluorescent protein or secreted placental alkaline phosphatase.
- 5 36. A method according to claim 34, wherein the heterologous gene comprises the gene for luciferase, chloramphenicol acetyl transferase,  $\beta$ -galactosidase, green fluorescent protein or secreted placental alkaline phosphatase.
37. An oligonucleotide sequence comprising a regulatory element of  
10 the nucleotide sequence  $\text{TTCN}_y\text{GAA}$ ,  $\text{TTAN}_y\text{GAA}$ ,  $\text{TTAN}_y\text{TAA}$ ,  $\text{TTTN}_y\text{GAA}$  and  $\text{TTTN}_y\text{TAA}$ , wherein N is independently selected from A, T, C or G and y is 2, 3, 4 or 5.
38. An oligonucleotide sequence according to claim 37 which is  
15 double stranded.
39. A selective oligonucleotide sequence comprising  
ANTTCNNNNGAANA (SEQ ID NO. 152) or TTTTCNNNNGAANT (SEQ ID NO. 153), wherein N is independently selected from A, T, C or G, and wherein the  
20 oligonucleotide sequence, when operably linked to a promoter and a heterologous gene, selectively binds to an activated transcriptional regulatory protein comprising a STAT protein activated by IL-4 and/or IL-13, and transcriptionally modulates the heterologous gene.

130

SUBSTITUTE SHEET (RULE 26)

WO 95/28482

PCT/US95/04477

131

40. A selective oligonucleotide according to claim 39 selected from the group consisting of ACTTCCCAAGAACA (SEQ ID NO. 154), TGTTCITGGGAAGT (SEQ ID NO. 155), ACTTCCCCGGAACA (SEQ ID NO. 156), TGTTCGGGGAAGT (SEQ ID NO. 157), ACTTCCCAGAACA (SEQ ID NO. 158), TGTTCGGGGAAGT (SEQ ID NO. 159), ACTTCCAGGAACA (SEQ ID NO. 160), TGTTCCTGGGAAGT (SEQ ID NO. 161), ACTTCCAAGAACA (SEQ ID NO. 162), TGTTCITGGGAAGT (SEQ ID NO. 163), ACTTCTTAAGAACA (SEQ ID NO. 164), and TGTTCCTTAAGAAGT (SEQ ID NO. 165), wherein the STAT protein comprises STAT6 protein.

10

131

SUBSTITUTE SHEET (RULE 26)

WO 95/28482

PCT/US95/04477

1/2

FIG. 1A

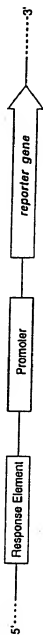
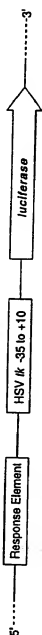


FIG. 1B

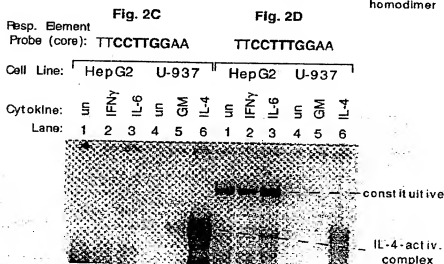
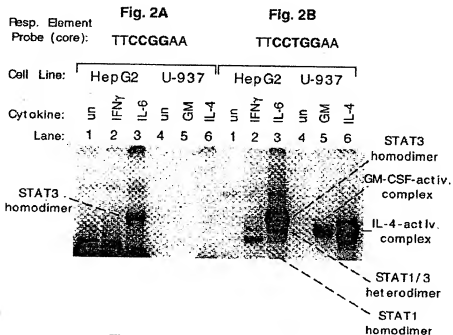


SUBSTITUTE SHEET (RULE 26)

WO 95/28482

PCT/US95/04477

2/2



SUBSTITUTE SHEET (RULE 26)

## INTERNATIONAL SEARCH REPORT

 International Application No.  
PCT/US 95/04477

|  |   |  |
|--|---|--|
| A. CLASSIFICATION OF SUBJECT MATTER  |   |  |
| IPC 6  | C12N15/11   | C12N15/19 C12N15/24 C12N15/85                      |
| According to International Patent Classification (IPC) or to both national classification and IPC  |   |  |
| B. FIELDS SEARCHED   |   |  |
| Minimum documentation searched (classification system followed by classification symbols)  |   |  |
| IPC 6 C12N   |   |  |
| Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  |   |  |
| Electronic data base consulted during the international search (name of data base and, where practical, search terms used)   |   |  |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT   |   |  |
| Category *   | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.                              |
| P, X   | WO-A-95 08801 (UNIV NEW YORK) 23 March 1995<br>see page 10; table I<br>seq. id. no. 2, 6, 9<br>---  | 1, 6, 18,<br>26, 38                                |
| Y  | MOLECULAR AND CELLULAR BIOLOGY,<br>vol. 14, no. 3, 1994<br>pages 1657-1668,<br>JUPING YUAN ET AL. 'THE SIGNALING<br>PATHWAYS OF INTERLEUKIN-6 AND GAMMA<br>INTERFERON CONVERGE BY THE ACTIVATION OF<br>DIFFERENT TRANSCRIPTION FACTORS WHICH BIND<br>TO COMMON RESPONSIVE DNA ELEMENTS'<br>see page 1662; figure 6<br>--- | 1-6,<br>8-25,<br>27-38                             |
| -/-  |   |  |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.  |   |  |
| * Special categories of cited documents:<br>"A" document defining the general state of the art which is not considered to be of particular relevance<br>"B" earlier document but published on or after the international filing date<br>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)<br>"O" document referring to an oral disclosure, use, exhibition or other means<br>"P" document published prior to the international filing date but later than the priority date claimed<br>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention<br>"X" document of particular relevance; the claimed invention cannot be understood without or cannot be considered to involve an inventive step when the document is taken alone<br>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, each combination being obvious to a person skilled in the art<br>"A" document member of the same patent family |   |  |
| Date of the actual completion of the international search  |   | Date of mailing of the international search report |
| 19 February 1996   |   | 01.03.96   |
| Name and mailing address of the ISA<br>European Patent Office, P.O. Box 5018 Patentamt 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Telex 31 631 epo nl,<br>Fax (+31-70) 340-3014   |   | Authorized officer<br><br>Halle, F                 |

Form PCT/ISA/210 (revised March 1997)



## INTERNATIONAL SEARCH REPORT

Int. Appl. No.  
PCT/US 95/04477

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |   |                                 |
|--|---|---------------------------------|
| Category *   | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.           |
| P, Y   | <p>BL000 ,<br/>vol. 83, no. 8, 15 April 1994<br/>pages 2063-2071,<br/>P. LAUB ET AL. 'RAPID ACTIVATION OF<br/>PROTEINS THAT INTERACT WITH THE INTERFERON<br/>GAMMA ACTIVATION SITE IN RESPONSE TO<br/>MULTIPLE CYTOKINES'<br/>see page 2069, right column, to page 2071,<br/>end of the text.<br/>cited in the application<br/>----</p> | <p>1-6,<br/>8-25,<br/>27-38</p> |
| A  | <p>EP-A-0 374 503 (TANIGUCHI TADATSUGU PROF<br/>DR) 27 June 1990<br/>see the whole document<br/>-----</p>   |                                 |

Form PCT/ISA/210 (continuation of standard form) (July 1992)

## INTERNATIONAL SEARCH REPORT

information on patent family members

Int. Appl. No.  
PCT/US 95/04477

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s)     | Publication<br>date  |
|---|---------------------|--------------------------------|----------------------|
| WO-A-9508001                              | 23-03-95            | AU-B- 7795094                  | 03-04-95             |
| EP-A-0374503                              | 27-06-90            | JP-A- 2273185<br>US-A- 5157115 | 07-11-90<br>20-10-92 |